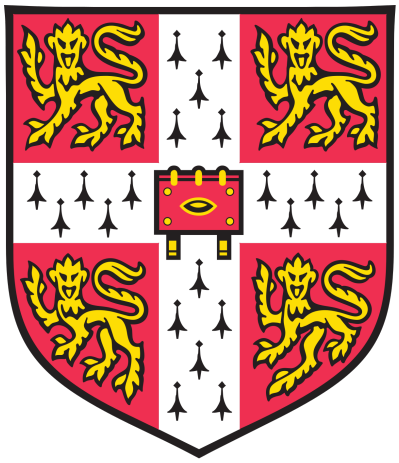
Developing mouse complex I as a model system: structure, function and implications for mitochondrial diseases

---



Fitzwilliam College

Ahmed-Noor Adam Agip

This dissertation submitted for the degree of *Doctor of Philosophy*

April 2018

Declaration

This dissertation, entitled "Developing mouse complex I as a model system: structure, function and implications for mitochondrial diseases", is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

This work was carried out between April 2014 and April 2018, under the supervision of Dr Judy Hirst. Information and data derived from other sources has been referenced and attributed accordingly.

Ahmed-Noor A. Agip
April 2018

Contributions to this thesis

Cryo-EM grid preparation, model building and image processing of the de-active mouse complex I (see Chapter 4) dataset were carried out by Dr. James Blaza from the MRC-MBU, University of Cambridge. As a result, we are joint first authors in a recently published research article (Ahmed-Noor A. Agip, James N. Blaza, Hannah R. Bridges, Carlo Viscomi, Shaun Rawson, Stephen P. Muench & Judy Hirst. (2018) Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states. Nature Structural & Molecular Biology *25*, 548-556).

Dedication

I solely dedicate this thesis to my dad and mum. In this world ‘dunya’, there are no two people who are more precious to me than my parents. They have dedicated their lives to our success (my siblings and I). They have showered us with their love, mercy and affection. For that, I love them abundantly. If I could, I would share my doctorate with them, and that still would not be enough to pay them back. To hooyo and aabo.

Acknowledgements

I would like to thank Judy for accepting me into her group and giving me the opportunity to work on excellent projects. I have improved tremendously under her guidance. I appreciate everything she has done for me and I wish her much success (i.e. Nobel prize).

To cope with the pressure that comes with the PhD, I had wonderful friends who supported me personally and scientifically. Febin was my first friend in the unit and was always there for me when I needed him. I am happy that you were comfortable with me as I was with you. I couldn't have asked for a better friend. Arnau is a wonderful friend and mentor, who always looked out for my best interest. There are few people like you in this world. You made my time more enjoyable in the lab. I miss both of them dearly. To Carlos, thank you for the support of the years and being such a magnanimous host when I visited you in Italy.

To my dear friend Chancie, thank you for the many years of guidance, 4 years and 6 months to be precise. Always thinking about others first. Your advice always weighs heavy on my scale. My only wish is for you to succeed.

I would also like to thank Carlo for mentoring me through years on the mice work that was essential to my PhD success and to the future direction of our lab. I couldn't have done my work without you.

I also appreciate other members of this unit, who over the years have taught me many techniques, such as Ricky and Andy on performing kinetic assays, and always being discrete when I messed up, to Erika for teaching me how to isolate mitochondria from mice. To the kind and wonderful Hannah Bridges, whom I have the utmost respect for, thank you for everything.

I also acknowledge Merlin and Andrew for setting up the computational clusters for cryo-EM image processing, to Mike, Shujing and Ian for mass spectrometry experiments and to Tobias for fruitful cryo-EM discussions.

To my secondary school friends, Daniel, Ahmed-Naajix, Askari, Chris, and Yasser, I appreciate you for keeping me humble.

And last but not least, to my siblings. My family is the most important thing to me in this world. You have looked after me since birth and for that I am grateful. I love you guys.

Abstract

Complex I (NADH:ubiquinone oxidoreductase), located in the mitochondrial inner membrane, is a major electron entry point to the respiratory chain. It couples the energy released from electron transfer (from NADH to ubiquinone) to the concomitant pumping of protons across the membrane, to generate an electrochemical proton motive force. Mammalian complex I is composed of 45 subunits, 14 of which comprise its simpler bacterial homologues. It is encoded by both the mitochondrial and nuclear genomes, and pathological mutations in both sets of subunits result in severe neuromuscular disorders such as Leigh syndrome. Several structures of mammalian complex I from various organisms have been determined, but the limited resolutions of the structures, which typically refer to poorly characterised enzyme states, has hampered detailed analyses of mechanistic features.

The first part of this thesis describes development of a method for purifying complex I from the genetically amenable and medically relevant model organism *Mus musculus* (mouse), in a pure, stable and active state. The enzyme from mouse heart mitochondria was then comprehensively characterised, to ensure the presence of all the expected subunits and co-factors, and to define its kinetic properties.

The second part of this thesis describes structural studies by single particle electron cryomicroscopy (cryo-EM) on the purified mouse enzyme in two distinct states, the 'active' and 'de-active' states. The active state was determined to 3.3 Å resolution, the highest resolution structure of a eukaryotic complex I so far. Subsequently, comparison of the two mouse structures, together with previously determined mammalian and bacterial structures, revealed variations in key structural elements in the membrane domain, which may be crucial for the catalytic mechanism. Moreover, in the high-resolution active mouse complex I structure a nucleotide co-factor was observed bound to the nucleoside kinase subunit NDUFA10.

Finally, complex I from the *Ndufs4* knockout mouse model, which recapitulates the effects of a human mutation that causes Leigh syndrome, was purified and subjected to kinetic and proteomic analyses. Following cross-linking and preliminary structural studies, it was concluded that the detrimental effects of deleting NDUFS4 are due to lack of stability of the mature complex.

Abbreviations

ADP	Adenosine diphosphate
$APAD^+$	3-acetylpyridine adenine dinucleotide
ATP	Adenosine triphosphate
BN-PAGE	Blue native-polyacrylamide gel electrophoresis
BS^3	bis(sulfosuccinimidyl)suberate
DDM	Dodecyl-maltoside
DMSO	Dimethyl sulphoxide
DQ	Decylubiquinone
EDTA	Ethylenediaminetetraacetic acid
EPR	Electron paramagnetic resonance
ETC	Electron transfer chain
FAD	Flavin adenine dinucleotide
FeCN	Potassium hexacyanoferrate (III) ($K_3Fe(CN)_6$)
FMN	Flavin mononucleotide
H_2O_2	Hydrogen peroxide
HAR	Hexaaminеruthenium ($Ru(NH3)_6^{3+}$)
HRP	Horseradish peroxidase
KCN	Potassium cyanide
LMNG	Lauryl Maltose Neopentyl Glycol
MALDI	Matrix-assisted laser desorption/ionization
mtDNA	Mitochondrial deoxyribonucleic acid
MWCO	Molecular weight cut-off

NAD^+	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide (reduced)
NBT	Nitroblue tetrazolium
ntDNA	Nuclear deoxyribonucleic acid
OD_{600}	Optical density at 600 nm
OXPHOS	Oxidative phosphorylation
PCR	Polymerase chain reaction
P_i	Inorganic phosphate
PMSF	Phenylmethanesulfonylfluoride
Q	Ubiquinone
QH_2	Ubiquinol
ROS	Reactive oxygen species
SDS-PAGE	Sodium dodecyl sulfate poly-acrylamide gel electrophoresis
SMP	Sub-mitochondrial particle
SOD	Superoxide dismutase
SQ	Semiquinone

Chapter 1: General introduction

1.1 Mitochondria

Mitochondria are often described as the "powerhouses of the cell" because they are the main source of adenosine triphosphate (ATP), which is the predominant energy currency in the cell. Besides their function in energy metabolism, they also have a plethora of other functions, including programmed cell death via the intrinsic pathway (Martinou and Youle, 2011), calcium homeostasis (Baughman et al., 2011; De Stefani et al., 2011), iron-sulphur cluster biogenesis (Braymer and Lill, 2017) and fatty acid biosynthesis (Hiltunen et al., 2010). Mitochondria generate ATP by coupling exergonic electron transfer through respiratory complexes I, II, III and IV, mediated by electron carriers such as cytochrome *c* and ubiquinone, to the endergonic translocation of protons by respiratory enzymes I, III and IV. This creates an electrochemical gradient, known as the proton motive force (PMF) that is harnessed by ATP synthase to drive the synthesis of ATP. This is known as oxidative phosphorylation and the enzymes that participate in this process are known as the oxidative phosphorylation enzymes.

Mitochondria are thought to have emerged from an endosymbiotic event between a facultative anaerobic α-proteobacterium and a host methanogenic archaebacterium (Embley and Martin, 2006), resulting in the first eukaryote. Although it is known that not all eukaryotes possess mitochondria (Karnkowska et al., 2016), they do contain mitochondria-like organelles, such as the hydrogenosome with its own subset of DNA (Boxma et al., 2005). This suggests that the acquisition of the ancestral mitochondrion was a prerequisite for the emergence of eukaryotes.

Mitochondria contain their own subset of DNA, the mitochondrial DNA (mtDNA) (Nass and Nass, 1963), which over the course of evolution was reduced through horizontal gene transfer into nuclear DNA (nDNA) (Mourier et al., 2001). Human mtDNA only encodes 13 proteins (directly involved in oxidative phosphorylation) and 24 RNAs (Anderson et al., 1981), with the vast majority of mitochondrial proteins encoded in the nDNA and imported into the mitochondria (Wiedemann et al., 2004). The mtDNA is transmitted from mother to offspring. Interestingly, in *Caenorhabditis elegans* embryos, paternal mitochondria are degraded by

autophagy (Sato and Sato, 2011), suggesting a functional relevance to the uniparental inheritance of mitochondria.

1.2 The architecture of mitochondria

Mitochondria are dynamic structures that are able to undergo fission and fusion depending on cellular stimuli, forming either highly dynamic interconnected networks or isolated mitochondria (Wai and Langer, 2016).

Mitochondria are made up of two compartments, an oxidising inter-membrane space and a reducing matrix, which are separated from the cytosol by a freely permeable outer-membrane and a selective permeable inner-membrane (Figure 1.1). The inner-membrane of mitochondria also forms invaginations called cristae, which greatly increases the surface area of the membrane. The oxidative phosphorylation enzymes are present on the cristae, with the respiratory enzymes on the flat surface of the cristae and the ATPase dimers on the ridges of the cristae bending the membrane (Kühlbrandt, 2015) (Figure 1.2).

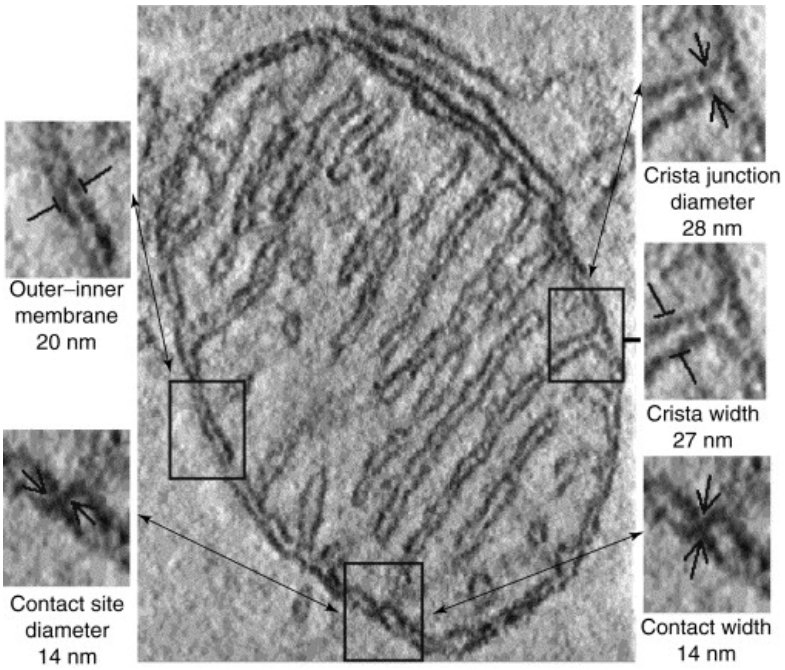


Figure 1.1: A cross-section of a *Gallus gallus* mitochondrion imaged using electron tomography. Various features are indicated in the enlarged images. This figure was taken from Frey and Mannella (Freya and Mannellab, 2000).

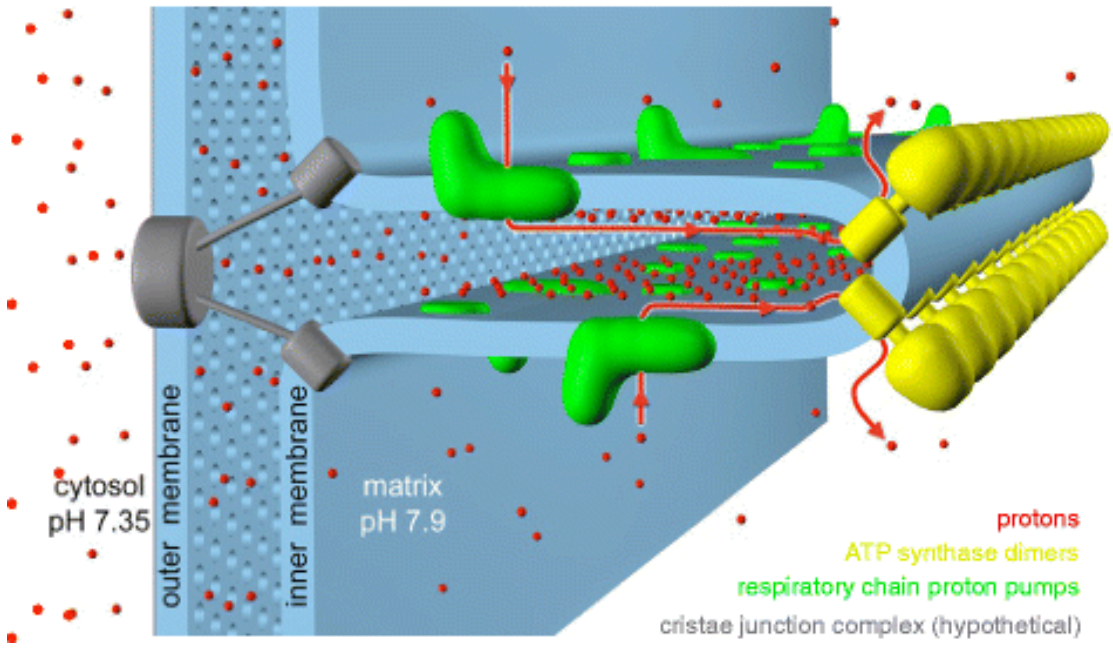


Figure 1.2: Organisation of protein complexes on cristae. Respiratory complexes (green) are shown in the flat region of the membrane whilst ATPase dimers (yellow) are shown on the ridge. The red arrow indicates the proton pathway through the enzymes and the cristae junction-forming protein is shown in grey. This Figure was taken from Kühlbrandt (Kühlbrandt, 2015).

1.3 Energy transduction in mitochondria

The majority of ATP in eukaryotic cells is produced by mitochondria (oxidative phosphorylation), with some synthesis occurring by glycolysis in the cytoplasm (substrate-level phosphorylation). Reducing equivalents, such as NADH, are generated via the citric acid cycle, which is fueled by pyruvate (from glycolysis), amino acids (amino acid degradation) and fatty acids (β-oxidation). Import of pyruvate, amino acids, and fatty acids (among other things) is facilitated by a class of membrane proteins called mitochondrial carriers (Kunji and Robinson, 2010) (Figure 1.3). Thereafter, the energy released from oxidizing electron transfer is used to drive oxidative phosphorylation (see Chapter 1.1)

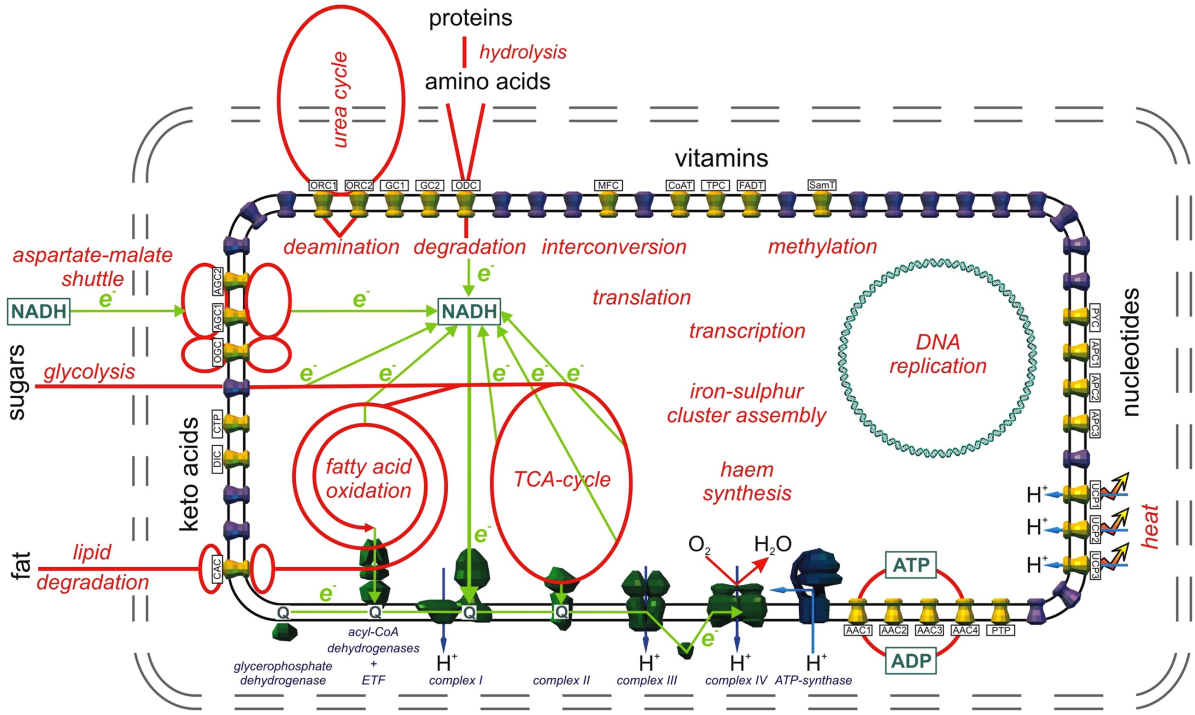


Figure 1.3: A general schematic of the processes that occur in mammalian mitochondria. The main metabolic pathways are shown in red. Known and unknown carrier proteins are shown in yellow and purple, respectively. Green arrows indicate electron flux, blue arrows indicate proton pumping and the curly red arrow indicates oxygen conversion to water. This Figure was taken from Kunji and Robinson (Kunji and Robinson, 2010).

There are many pathways of delivering electrons to the ubiquinone pool, such as the one through the electron transfer flavoprotein:ubiquinone oxidoreductases (Watmough and Frerman, 2010), but the two main electron flux pathways through the respiratory chain that result in generation of the proton-motive force that is utilised to produce ATP are the NADH to O_2 pathway (catalysis complexes I, III and IV) and the succinate to O_2 pathway (catalysis by complexes II, III and IV) (Figure 1.4). A brief summary of the reactions that occurs in each complex is given below.

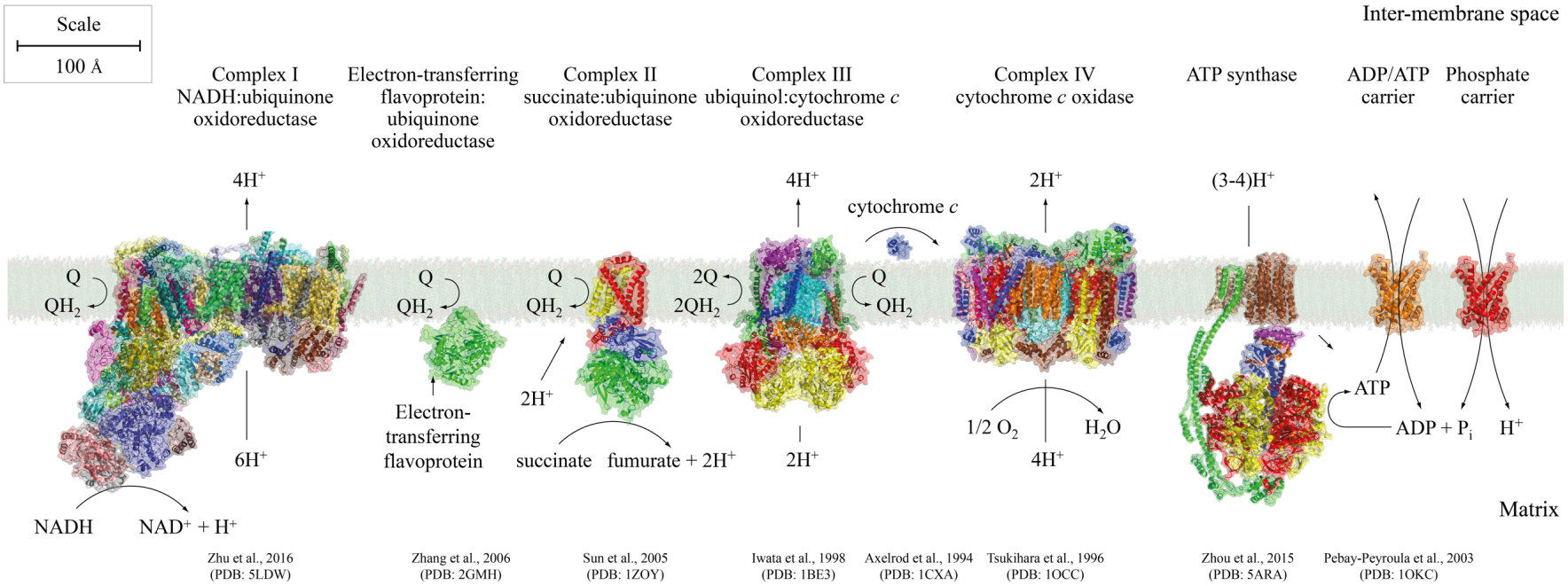


Figure 1.4: The major protein complexes involved in mitochondrial energy transduction. For clarity, other proteins of the mitochondrial inner membrane are not shown. Images of three-dimensional structures and electron density maps were generated using PyMOL (DeLano scientific, 2002). Complexes I (PDB: 5LDW), III (PDB: 1BE3) and IV (PDB: 1OCC) are from *B. taurus* (Zhu et al., 2016; Iwata et al., 1998; Tsukihara et al., 1996). The structures of the electron-transferring flavoprotein (PDB: 2GMH) and complex II (PDB: 1ZOY) are from porcine mitochondria (Zhang et al., 2006; Sun et al., 2005). Cytochrome *c* (PDB: 1CXA) is from *Rhodobacter*

NADH:ubiquinone oxidoreductase (complex I): Complex I links the oxidation of NADH to NAD^+ to the reduction ubiquinone to ubiquinol. The energy released from the redox chemistry is used to pump four protons across the membrane to maintain a PMF of ~180 mV (Jones et al., 2017). The overall reaction is given below.

$$NADH + ubiquinone + 2H^+_{matrix} + 4H^+_{matrix} \leftrightarrow NAD^+ + H^+_{matrix} + ubiquinol + 4H^+_{inter\text{-}membrane\ space}$$

Succinate:ubiquinone dehydrogenase (complex II): Complex II is the only enzyme in the respiratory chain involved in the citric acid cycle. The flavin adenine dinucleotide (FAD) co-factor of the enzyme oxidises succinate to fumarate on the matrix side. The electrons are then transferred via a chain of three iron-sulphur clusters to reduce ubiquinone to ubiquinol. The reaction is not electrogenic, with no transfer of charge across the inner-membrane, and therefore dose not contribute to the PMF generation. The overall reaction is given below (also see Figure 1.5).

$$succinate + ubiquinone \leftrightarrow fumarate^+ + ubiquinol$$

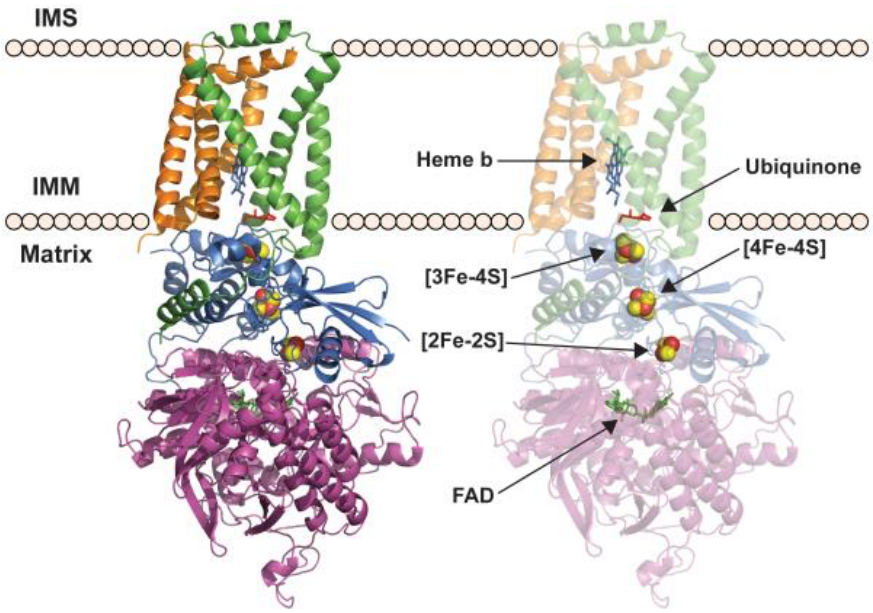


Figure 1.5: The structure porcine succinate dehydrogenase (complex II). Electron transfer occurs from FAD to ubiquinone via three iron-sulphur clusters (all indicated in the structure). IMM and IMS denote inner mitochondrial membrane and inner membrane space, respectively. This figure was taken from Van Vranken *et al.*, (Van Vranken et al., 2015).

Ubiquinol:cytochrome *c* oxidoreductase (complex III): Complex III operates by a Q-cycle mechanism. An ubiquinol binds near the inter-membrane space (Q_p site) and passes one electron to the Rieske protein, which then passes it to cytochrome *c* (for the complex IV reaction), releasing two protons into the inter-membrane space. The second electron is transferred through two haems to a bound ubiquinone near the matrix side (Q_n site), reducing it to a semi-ubiquinone. The oxidation of ubiquinol to ubiquinone at the Q_p binding site is repeated, with an electron transferred to a second cytochrome *c* through the Rieske protein (releasing two protons into the inter-membrane space), and the other electron being transferred to the semi-ubiquinone at the Q_n site, forming ubiquinol by up taking two protons from the matrix side. The overall reaction is given below (also see figure 1.6).

$$\text{ubiquinone} + 2\ \text{ubiquinol} + 2H^{+}_{\text{matrix}} + 2\ \text{cytochrome}\ c_{ox} \leftrightarrow$$
$$2\ \text{ubiquinone} + \text{ubiquinol} + 4H^{+}_{\text{inter-membrane space}} + 2\ \text{cytochrome}\ c_{red}$$

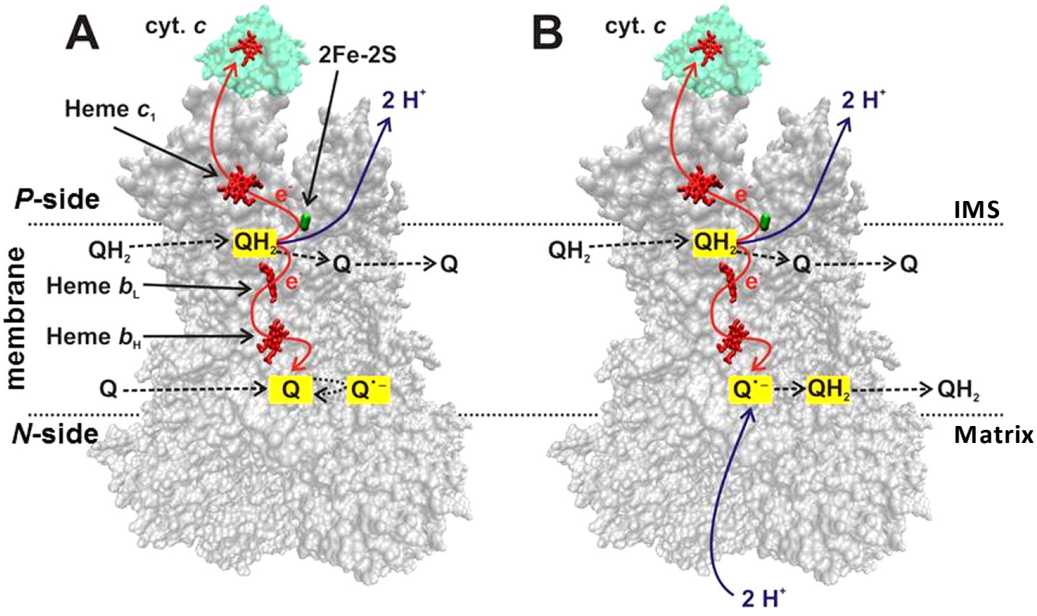


Figure 1.6: Q-cycle reaction scheme of Ubiquinol:cytochrome c oxidoreductase complex III. The figure illustrates the successive oxidation of two ubiquinols (A and B), resulting in the reduction of ubiquinone and cytochrome *c*. The proton pathways are indicated with the blue arrow. P-side and N-side denote the inter-membrane space (IMS) and the matrix, respectively. The electron transfer pathways through the co-factors are shown in red. This figure was taken from Wikström *et al.,* (Wikström et al., 2015)

Cytochrome *c* oxidase (complex IV): Electrons are transferred, one at a time from cytochrome c_{red} (reduced by complex III) to copper centre Cu_A, then haem *a* and finally to haem-a_3-Cu_B centre-haem (Figure 1.7). Consequently, molecular oxygen is reduced to two H_2O in each catalytic cycle, which requires four protons and four electrons (from four cytochrome c_{red} molecules). In the process, four protons are translocated across the membrane for every molecule of oxygen reduced (Figure 1.7).

$$4 \text{ cytochrome } c_{red} + 8H^+_{matrix} + O_2 \leftrightarrow$$
$$4 \text{ cytochrome } c_{ox} + 4H^+_{inter\text{-}membrane\ space} + 2H_2O$$

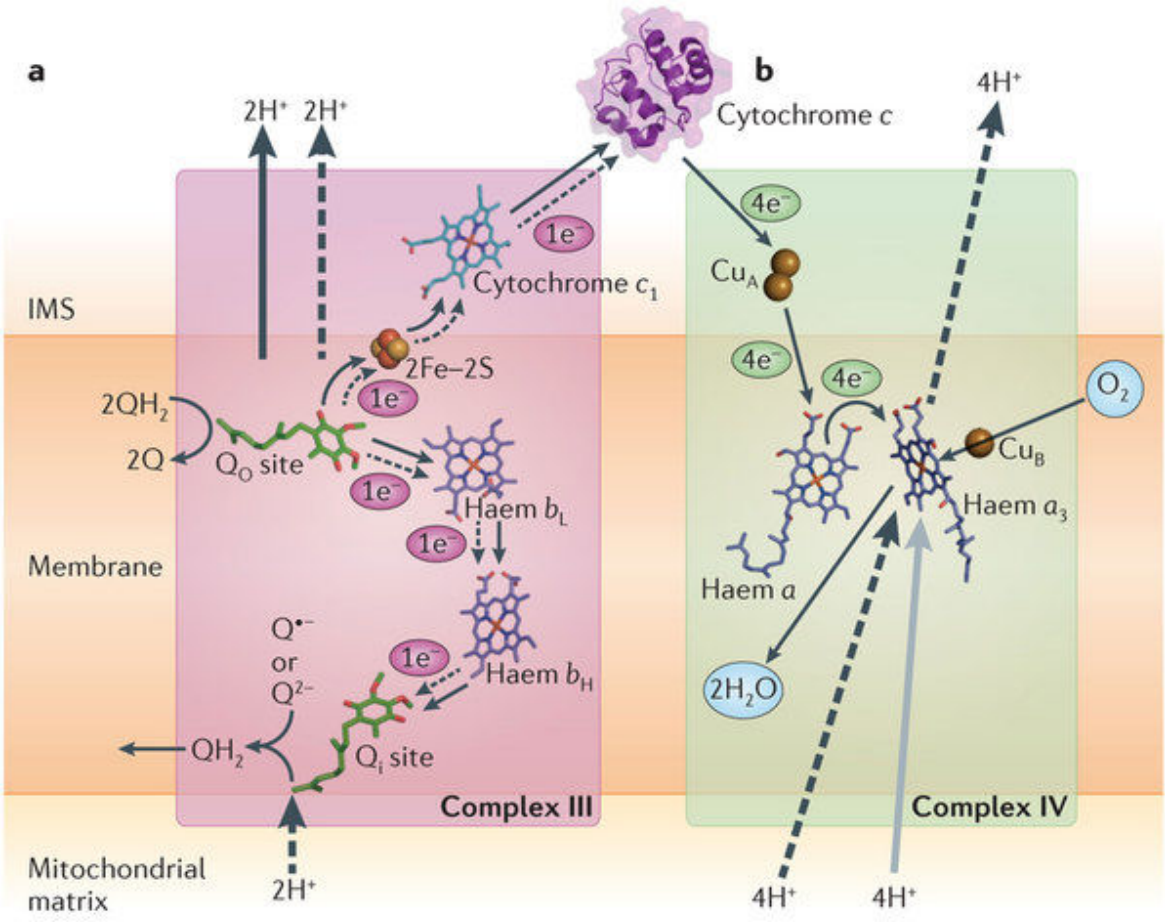


Figure 1.7: Reaction scheme of both complex III and IV. Electron transfer, catalysis and proton translocation for complex III (a) and complex IV (b) are shown. This figure was taken from Sazanov (Sazanov, 2015)

Overall, the favourable electron transfer from NADH to ½ O_2 (catalysis by complexes I, III and IV) provides the energy to drive 10 H^+ against the proton-motive force into the inter-membrane space.

ATP synthase (complex V): The proton-motive force generated by the respiratory chain complexes is utilised by ATP synthase to produce ATP from ADP and inorganic phosphate (P_i). The c subunits in the complex (8 are present in the mammalian enzyme) which makes up the turbine-like structure contain glutamate residues which are protonated through a half-channel in subunit a on the inter-membrane space (Figure 1.8). The proton motive force induces a clock-wise rotation (looking down on the c-ring from the inter-membrane space) of the c-ring. As rotation occurs, protons are transferred by glutamate residues from the inter-membrane space to the matrix (Nakamoto et al., 2008).

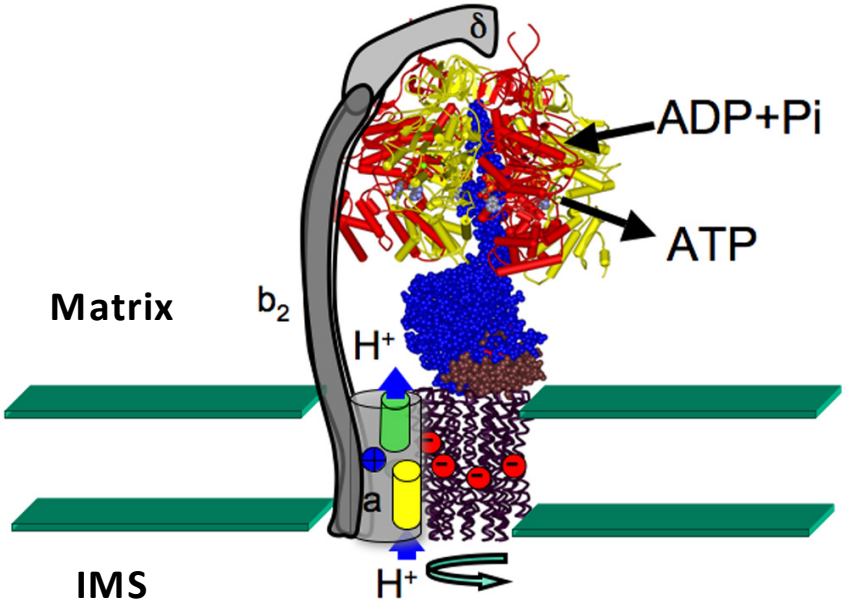


Figure 1.8: Production of ATP by ATP synthase. A simple scheme of one of the reactions carried out by ATP synthase. The red spheres are glutamates of the *c*-ring and the yellow and green cylinders represent the half-channels in the *a* subunit where protons are translocated from the inter-membrane space (IMS) to the matrix side during ATP synthesis. This figure was taken from Nakamoto *et al.,* (Nakamoto et al., 2008)

The rotation of the c-ring causes a like-wise rotation of the central stalk (subunit γ), connecting the proton-translocating (F_o) domain to the catalytic (F_1) domain. F_1 is made up of three domains, each containing an α and a β subunit, with each domain able to catalyse ATP

production. The rotation of subunit γ causes conformational changes in the catalytic F_1 domain (Figure 1.9). The F_1-associated catalytic subunit β is initially in an open confirmation (β_E) which allows binding of ADP and inorganic phosphate. Following the rotation of subunit γ, subunit β changes to a loose (β_{DP}) and then a tight (β_{TP}) conformation, catalysing ATP production (Figure 1.9). For every complete rotation of the c-ring, 8 protons are taken up from the inter-membrane space and released in the matrix, resulting in the generation of 3 ATP molecules. This means that for every 10 H^+ pumped into the inter-membrane space by the respiratory chain complexes (NADH to ½ O_2), complex V generates ~3.8 molecules of ATP This does not take into account the proton required for phosphate transfer into mitochondria.

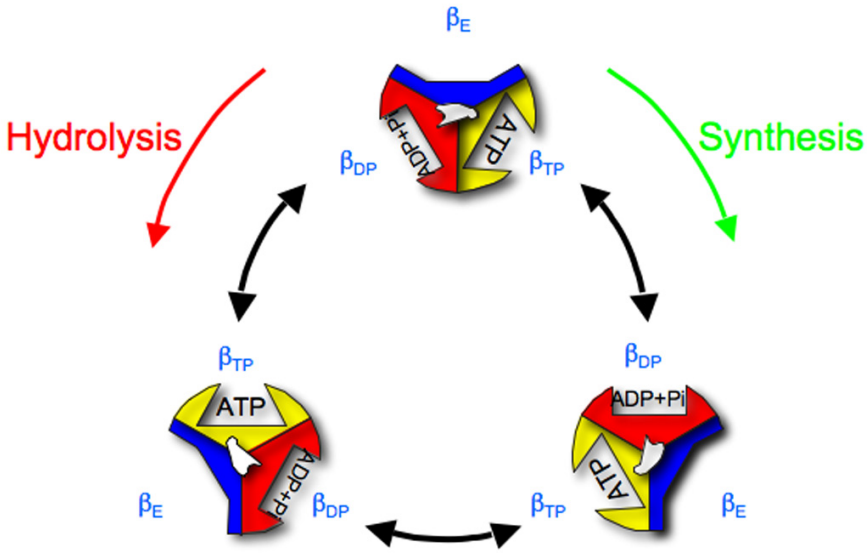


Figure 1.9: ATP production by the F_1 domain of ATP synthase. Conformational changes in the β subunits result in either synthesis of ATP or its hydrolysis (dependent on the rotational direction of subunit γ. This figure was taken from Nakamoto *et al.,* (Nakamoto et al., 2008).

1.4 Mammalian complex I

Mammalian complex I is a large integral membrane protein (~960 kDa) made up of 45 subunits, of which 44 are unique (Vinothkumar et al., 2014). Mammalian complex I shares 14 conserved core subunits with its bacterial counterparts, which is about half its mass. These shared core subunits constitute the minimal set of proteins required for catalysis by complex I. There are differences in the complex I subunit nomenclature between different species (Table 1.1). The human complex I subunit nomenclature will be used from here on for simplicity, regardless of the species referenced.

Mammals		Yeast	Bacteria	
B. taurus	***H. sapiens***	***Y. lipolytica***	***E. coli***	***T. thermophilus***
75 kDa	NDUFS1	NUAM	NuoG	Nqo3
51 kDa	NDUFV1	NUBM	NuoF	Nqo1
49 kDa	NDUFS2	NUCM	NuoD*	Nqo4
30 kDa	NDUFS3	NUGM	NuoC*	Nqo5
24 kDa	NDUFV2	NUHM	NuoE	Nqo2
PSST	NDUFS7	NUKM	NuoB	Nqo6
TYKY	NDUFS8	NUIM	NuoI	Nqo9
ND1	ND1	NU1M	NuoH	Nqo8
ND2	ND2	NU2M	NuoN	Nqo14
ND3	ND3	NU3M	NuoA	Nqo7
ND4	ND4	NU4M	NuoM	Nqo13
ND5	ND5	NU5M	NuoL	Nqo12
ND6	ND6	NU6M	NuoJ	Nqo10
ND4L	ND4L	NULM	NuoK	Nqo11
42 kDa	NDUFA10			
39 kDa	NDUFA9	NUEM		
18 kDa	NDUFS4	NUYM		
15 kDa	NDUFS5	NIPM		
13 kDa	NDUFS6	NUMM		
10 kDa	NDUFV3			
AGGG	NDUFB2			
ASHI	NDUFB6	NIAM		
ESSS	NDUFB11	NESM		
KFYI	NDUFC1			
MNLL	NDUFB1			
MWFE	NDUFA1	NIMM		
PDSW	NDUFB10	NIDM		
PGIV	NDUFA8	NUPM		
SDAP-α	NDUFAB1	ACPM		
SDAP-β	NDUFAB1	ACPM		
SGDH	NDUFB5			
B22	NDUFB9	NI2M		
B18	NDUFB7	NB8M		
B17.2	NDUFA12	N7BM		
B17	NDUFB6			
B16.6	NDUFA13	NB6M		
B15	NDUFB4	NB5M		
B14.7	NDUFA11	NUJM		
B14.5a	NDUFA7	NUZM		
B14.5b	NDUFC2			
B14	NDUFA6	NB4M		
B13	NDUFA5	NUFM		
B12	NDUFB3	NB2M		
B9	NDUFA3	NI9M		
B8	NDUFA2	NI8M		
		NUXM		
		NEBM		
		NUNM		
		NUUM		

Table 1.1: Complex I subunit nomenclature from different species. This table was kindly provided by Dr. Febin Varghese.

1.5 General structure of complex I

Since the first purification of eukaryotic complex I 56 years ago (Hatefi et al., 1962), many attempts have been made to solve its structure. The first success came with negatively stained complex I from *Neurospora crassa*, in which its distinguishing 'L' shape feature was first observed (Guénebaut et al., 1997; Hofhaus et al., 1991). The structure of complex I was shown to be comprised of two major domains, a hydrophobic domain that is embedded in the membrane and a hydrophilic domain that is protruding orthogonally from the membrane domain and into the matrix/cytoplasm. Although the general shape of complex I was determined from the structure, the resolution of the reconstructions was too low (~30 Å) to be able to observe useful structural information (i.e. subunits location, secondary structure features.

1.6 The structure of bacterial complex I

The first high resolution structure for complex I, obtained by X-ray crystallography, was that of the hydrophilic domain of complex I from *Thermus thermophilus*, first at 4.0 Å and then at 3.3 Å (Hinchliffe and Sazanov, 2005; Sazanov and Hinchliffe, 2006). The hydrophilic arm consists of an N-module (the site of NADH oxidation), with subunits NDUFV1, NDUFV2), NDUFS1, and the Q-module (the site of quinone reduction), with subunits NDUFS2, NDUFS3, NDUFS7 and NDUFS8 (Figure 1.10). In addition, there was an additional supernumerary subunit present in the structure, NQO15, which is exclusive to *T. thermophilus*.

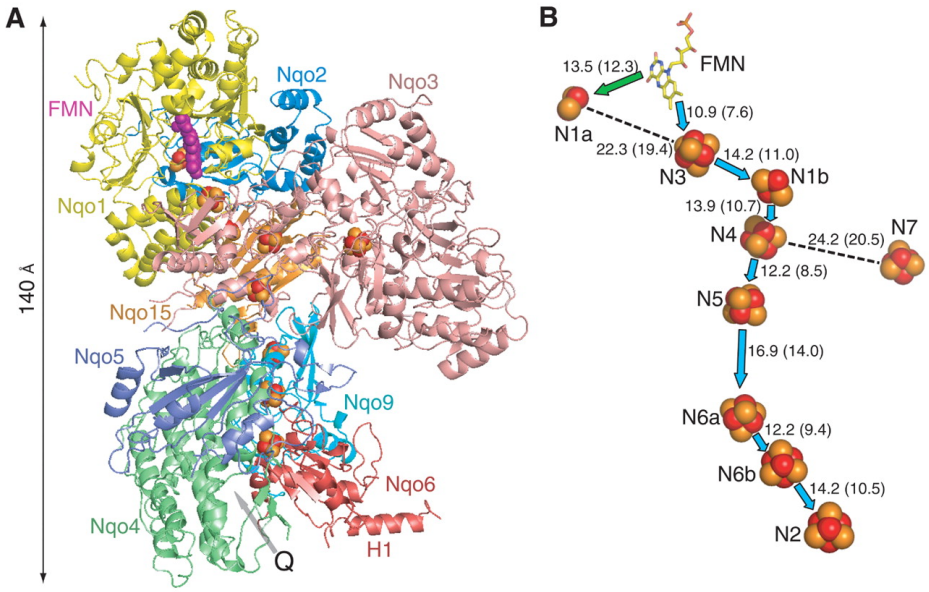


Figure 1.10: Architecture of the hydrophilic domain of complex I from *T. thermophilus*: (A) A side view of the hydrophilic subunits of complex I (cartoon representation), iron-sulphur clusters (red and yellow sphere representation) and FMN (magenta sphere representation). Q denotes the quinone-binding site. (B) The main electron transfer pathway (arrow) and possible electron transfer pathway (green), with centre-to-centre (and edge-to-edge in parenthesis) distances in Å. See table 1.1 for complex I subunit nomenclature. This figure was taken from Sazanov and Hinchliffe (Sazanov and Hinchliffe, 2006).

The structure shows that all of the catalytically-relevant co-factors, the FMN, and the iron-sulphur clusters are surrounded by the core subunits of the hydrophilic arm. For example, the N-terminal domain of NDUFV1 contains a Rossman-like fold, enabling it to bind the FMN, and the N-terminal [FeFe] domain of NDUFS1 binds three iron-sulphur clusters. There are nine iron-sulphur clusters (eight in eukaryotes) present in the structure of the *T. thermophilus* hydrophilic arm (Figure 1.10). Seven of the clusters form a chain from the FMN to the base of the hydrophilic arm where quinone binds, and they are all within favourable electron transfer distance from each other (within 14 Å) (Page et al., 2003). This provides a pathway of electron transfer from NADH (via the FMN) to quinone (via N2, the terminal cluster in the chain).

The structure of the membrane domain of complex I, which is the site of proton pumping, was solved by X-ray crystallography, at 3.0 Å from *Escherichia coli* (Efremov and Sazanov, 2011). The structure contained six out of the seven core subunits of complex I present in the membrane domain, ND5, ND4, ND2, ND4L, ND6 and ND3 (Figure 1.11). The structure lacked ND1

because it dissociated during the protein purification. Remarkably, the membrane domain contains no co-factors and is too far away from the site of ubiquinone reduction to be coupled directly, therefore an indirect mechanism for coupling ubiquinone reduction by NADH to proton pumping in the membrane domain has to occur.

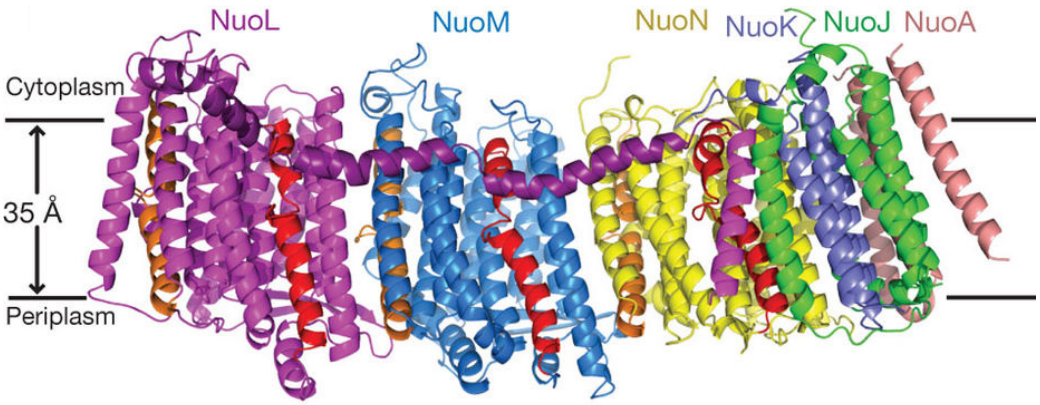


Figure 1.11: Structure of the membrane domain of complex I from *E. coli*. A cartoon representation of the assigned membrane subunits (labelled), with the predicted span of the lipid bilayer indicated. See table 1.1 for complex I subunit nomenclature. This figure was taken from Efremov and Sazanov (Efremov and Sazanov, 2011).

Subunits ND5, ND4 and ND2 are homologous to Na^+/H^+ Mrp antiporters (Mathiesen and Hägerhäll, 2002), which suggests that they are involved in proton pumping. The antiporter-like subunits are divided into two halves, related to each other by inverted symmetry along the membrane plane. Each symmetry-related half contributes a half-channel, in which one half has access to the inter-membrane/periplasmic space while the other half has access to the matrix/cytoplasm. The two half-channels are disconnected from one another but are linked by charged residues and probably water molecules to mediate proton transfer. Each half of the antiporter-like subunits possess a discontinuous helix, which is a helix disrupted at its centre by a loop (Figure 1.12). This confers flexibility on the helix that may allow it to undergo conformational changes during proton pumping (Efremov and Sazanov, 2011). On the discontinuous helices are highly conserved charged residues that lie in the middle of the membrane, making them ideal candidates for coupling electron transfer to proton pumping. In addition, ND5 contains a long C-terminal transverse helix that spans most of the membrane domain. The residues of the transverse helix are poorly conserved and while its function is not known, it may provide a means of synchronizing proton pumping or improving membrane domain stability. Together with the antiporter-like subunits, ND4L and ND6 are said to

contribute the fourth distinct proton channel, mediated by a path of conserved charged residues, although the exact pathways through these subunits are unclear (Figure 1.13). Although the number of channels is consistent with complex I pumping four protons across the membrane (Jones et al., 2017), visualisation of water channels which conduct proton pumping has not yet been possible.

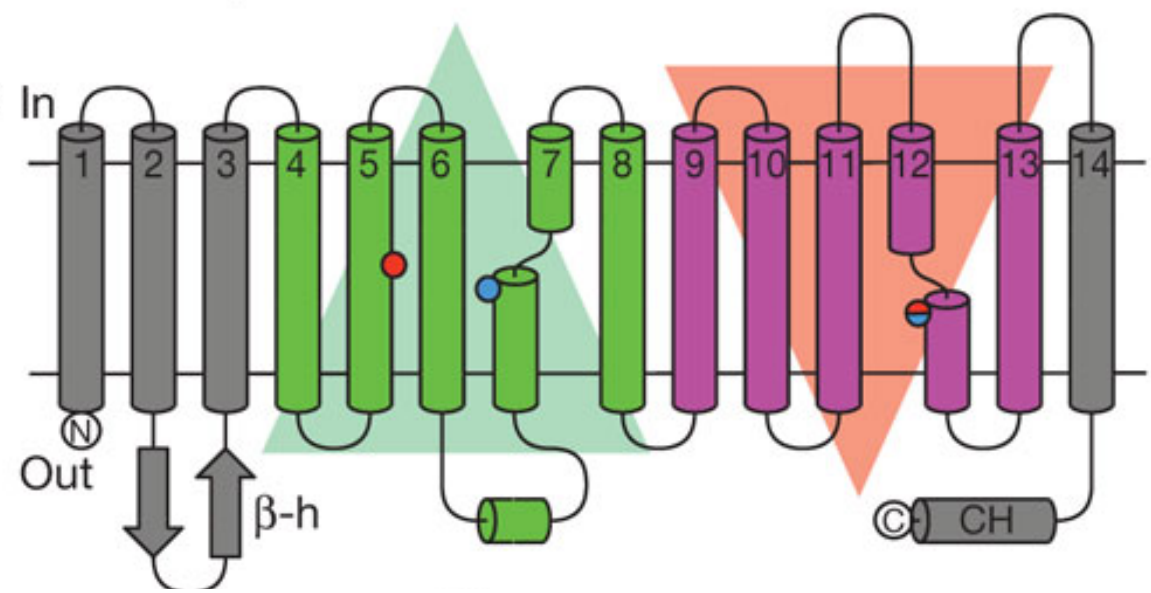


Figure 1.12: Topology of the antiporter-like subunits ND2, ND4 and ND5. The two symmetry-related halves are shown in green and magenta. Critical charged residues are shown as red or blue spheres. This figure was taken from Efremov and Sazanov (Efremov and Sazanov, 2011).

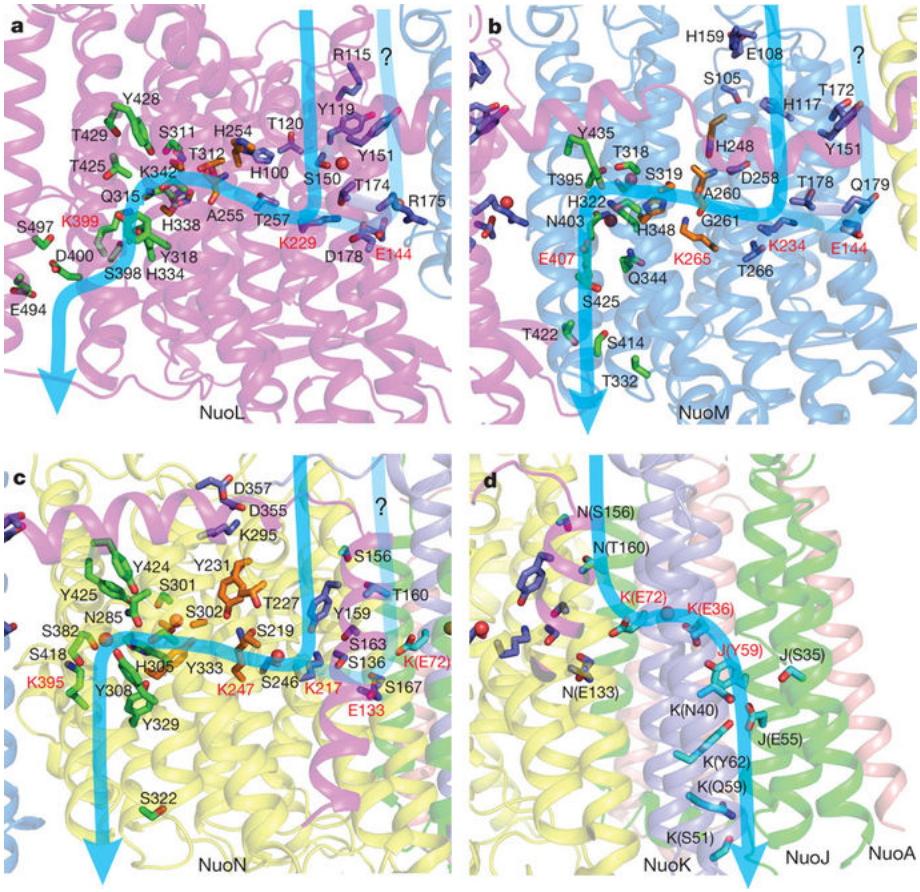


Figure 1.13: Putative proton channels in the membrane domain. Essential charged residues predicted to be involved in proton translocation are labelled red. Predicted water molecules are shown as red spheres, with the path of proton translocation indicted by blue arrow. This figure was taken from Efremov and Sazanov (Efremov and Sazanov, 2011).

The first crystal structure of the entire respiratory complex I at 3.3 Å was solved from *T. thermophilus* (Baradaran et al., 2013). The structure contained all of the core subunits from the membrane and hydrophilic domains (Figure 1.14a). The arrangement of the antiporter-like subunits in the *T. thermophilus* structure is similar to that in the *E. coli* structure. The first half-channel (cytoplasmic side) subunits have a lysine residue on the discontinuous helix and a nearby interacting glutamate residue (Figure 1.14b). The second half-channel of the antiporter-like subunits have either a lysine or a glutamate at the discontinuous helix. The two half channels are connected by either a histidine or lysine.

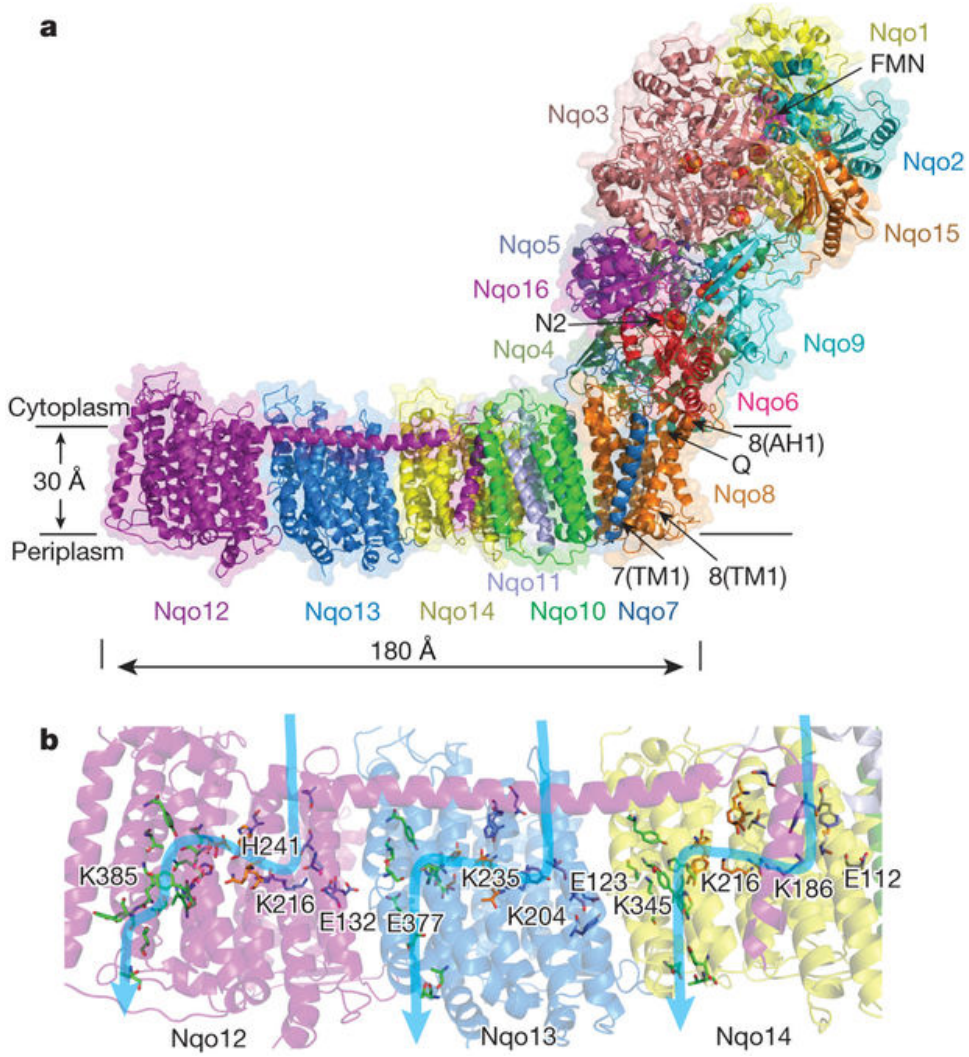


Figure 1.14: Structural overview of complex I from *T. thermophilus*. (a) A cartoon representation of the entire complex I, with the predicted span of the lipid bilayer indicated. (b) Predicted channels (blue arrow) by conserved charged residues (labelled). See table 1.1 for complex I subunit nomenclature. This figure was taken from Baradaran et al., (Baradaran et al., 2013).

A novel feature of the structure is that the predicted ubiquinone-binding chamber, which is long and narrow, extends 15 Å above the membrane surface (Baradaran et al., 2013). This was partially confirmed by soaking the *T. thermophilus* complex I crystals in piericidin, which is a quinone analogue and potent complex I inhbitor. This long channel means that quinone has to 'diffuse' into complex I to become reduced by electrons from the N2 iron-sulphur cluster. Sazanov and co-workers postulated that long-range conformational changes occur by electrostatic repulsion of key charged residues across the membrane plane (Figure 1.14b) upon quinone reduction, but this has yet to be experimentally shown.

1.7 General mechanism of complex I

In order for electron transfer to occur in complex I, NADH must first bind in the binding pocket of NDUFV1, which catalyses the hydride transfer from NADH to the FMN (Birrell and Hirst, 2013). The electrons from the reduced FMN are transferred, one at a time through a series of iron sulphur clusters (Figure 1.15), with the pathway (shown in figure 1.10) being NADH → FMN → N3 → N1b → N4 →N5 → N6a → N6b → N2 → Q, where N and Q denote the clusters and quinone, respectively. The transfer of electrons is not rate limiting for catalysis, with two electrons sequentially delivered to the terminal N2 cluster in only ~90 μs in *E. coli* complex I (Verkhovskaya et al., 2008).

Following reduction of the clusters, two electrons are sequentially transferred to quinone, with its head group ~15 Å above the membrane plane, as inferred from an analogous structure bound to the *T. thermophilus* complex I (Baradaran et al., 2013). Coupled to its reduction, the quinone is protonated by nearby tyrosine and histidine residues (note: this is predicted and not experimentally determined) (Sharma et al., 2015) that are within hydrogen-bonding distance to the quinone head-group.

Electron transfer by the hydrophilic arm of complex I is coupled to the transfer of four protons in the membrane arm of the complex I (Jones et al., 2017), most likely through the antiporter like subunits (see figure 1.13). How exactly electron transfer is energetically coupled to proton pumping has yet to be determined, but a series of charged residues in the central axis of the hydrophobic membrane (i.e. in the antiporter-like subunits) are predicted to be involved in coupling the reduction of quinone to the pumping of protons, by inducing conformational changes in the membrane domain (Figure 1.15) (Efremov and Sazanov, 2011). Further work is required to elucidate the coupling mechanism of complex I.

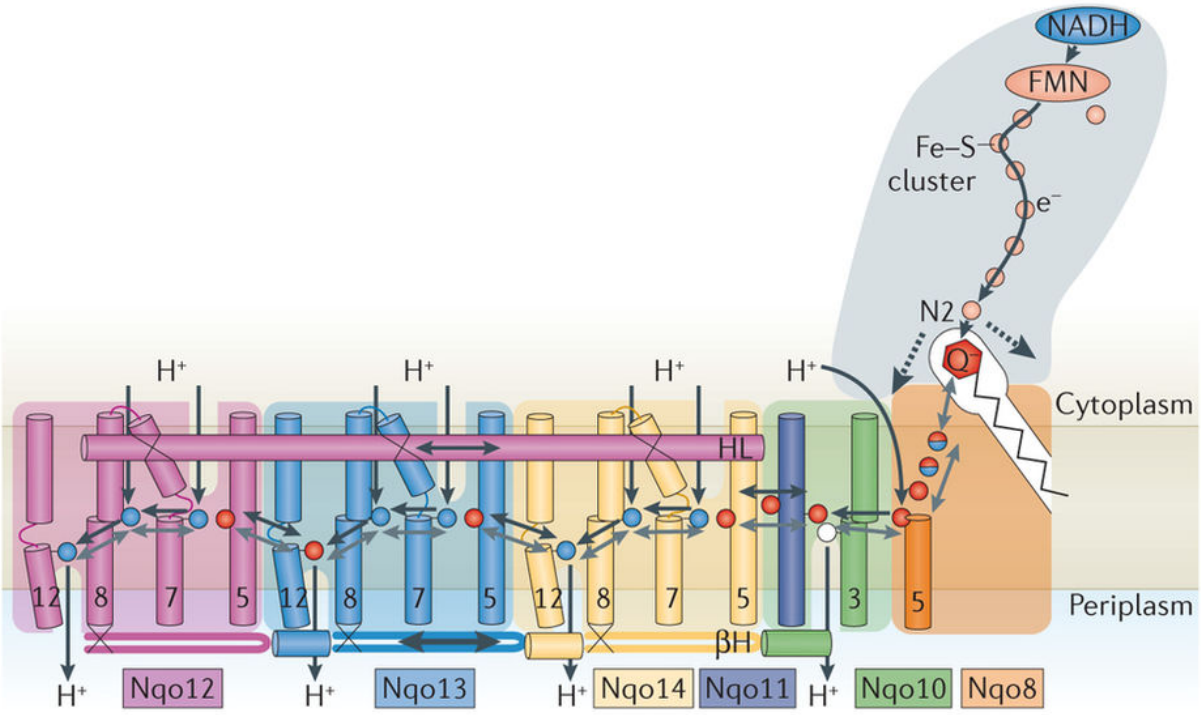


Figure 1.15: Proposed coupling mechanism of complex I. Electron transfer and the subsequent proton transfer are shown. Charged residues involved in energy coupling are shown as spheres, red for acidic and blue for basic residues. See table 1.1 for complex I subunit nomenclature. This figure was taken from Sazanov (Sazanov, 2015).

1.8 Active to de-active transition of mitochondrial complex I

Mitochondrial complex I undergoes a transition from an active state to a de-active state. This was first observed in horse heart preparations, which exhibited bi-phasic NADH oxidation kinetics, characterised by a lag phase transitioning to a more rapid steady-state phase (Minakami et al., 1964). It was later shown that the active to the de-active transition of complex I is dependent on a lack of substrate and has a temperature dependence (i.e. incubation at 30 °C causes deactivation of the enzyme), as observed in bovine heart sub-mitochondrial particles (SMPs) (Kotlyar and Vinogradov, 1990). They also found the opposite to be true, where incubation of bovine heart SMPs with NADH resulted in the activation of the enzyme (i.e. no lag phase present in the kinetic measurements).

Proteomic analyses and selective labelling of both active and de-active bovine heart SMPs showed that a highly conserved cysteine residue on the ND3 subunit (cysteine 39 in mammals) is a diagnostic for the active to the de-active transition of complex I (Galkin et al., 2008). The cysteine residue is located in a loop between transmembrane helix 1 and 2 of ND3, which in the active state is observed to bind in a cavity near the NDUFS2, NDUFS7 and ND1 subunits (Figure 1.16). The cysteine-containing loop is believed to become disordered and exposed to the matrix when complex I is de-activated (i.e. incubation of mitochondrial membranes at 30 °C) and the cysteine becomes susceptible to *N*-ethylmaleimide (NEM) modification, consistent with structural and biochemical data (Blaza et al., 2018; Galkin et al., 2008; Zhu et al., 2016).

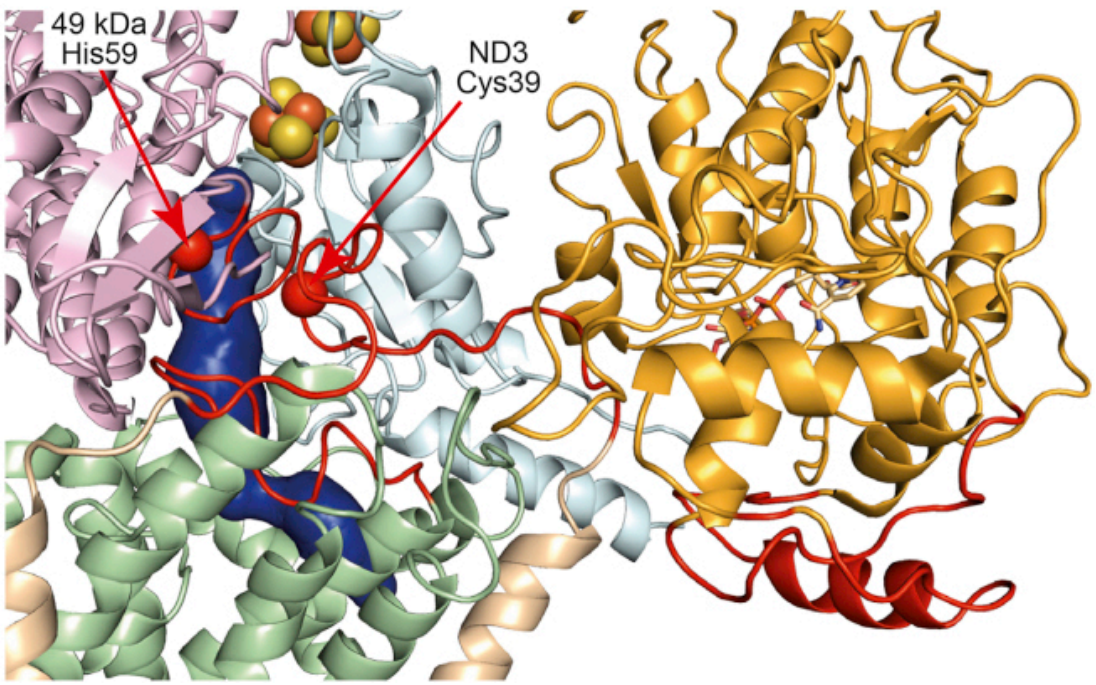


Figure 1.16: Location of cysteine 39 on ND3 loop. Highly conserved cysteine 39 and histidine 59 (red spheres) are labelled. Disordered regions in the de-active bovine complex I structure (5LDW) and the ubiquinone binding channel are shown in red and blue, respectively (Zhu et al., 2016). ND1 (green), ND3 (beige), NDUFS7 (light blue), NDUFS2 (pink) and NDUFA9 (gold) are shown in cartoon representation. This figure was taken from Blaza *et al.*, (Blaza et al., 2018).

There are pathophysiological consequences to the deactivation of complex I in ischemia-reperfusion injury (Chouchani et al., 2014). During ischemia, it is thought that a lack of oxygen to complex IV stalls electron transfer through the respiratory chain, resulting in the de-

activation of complex I (by a lack of substrate turnover) and a buildup of metabolites, such as succinate (Chouchani et al., 2014). The re-introduction of oxygen (reperfusion) causes complex I to reduce oxygen to superoxide via reverse electron transfer (opposite of the scheme shown in Figure 1.15), which is driven by the proton motive force and correlated to succinate buildup (Chouchani et al., 2014). It was shown in mouse heart that the injury caused by a surge in reactive oxygen, can be protected against by some extent by reversible S-nitrosation of cysteine 39 of complex I (Chouchani et al., 2013).

It is known that lack of substrates supposedly drives complex I into the de-active state (Kotlyar and Vinogradov, 1990), however, the exact mechanism and utility of deactivation is not known, nor is the *in vivo* relevance. Furthermore, bacterial complex I exists only in the active form, including *Paracoccus denitrificans*, which has the conserved ND3 cysteine (unlike *E. coli* and *T. thermophilus*) (Maklashina et al., 2003), the reason for this is not known.

1.9 The structure of yeast complex I

The first density map to reveal secondary structure features of eukaryotic complex I was that of *Yarrowia lipolytica* at 6.3 Å, using X-ray crystallography. The iron-sulphur clusters are apparent, with a similar arrangement to the *T. thermophilus* complex I (Hunte et al., 2010). However, because of its low resolution, no structural model could be built for the fungal enzyme. Thereafter, a crystal structure of complex I from the same organism was obtained with a resolution of 3.6 Å (Zickermann et al., 2015). Although at a higher resolution, only one supernumerary subunit was modelled in (NDUFA9) (Figure 1.17). However, 82 transmembrane helices have been modelled, with only the transmembrane helices for the core subunits (64 helices) having been assigned to their respective subunits. The arrangement of core subunits is globally identical to that of previously solved complex I structures (Figure 1.17).

The enzyme was co-crystallised with 2-decyl-4-quinazolinylamine (DQA), which like piericidin, is a quinone analogue. A strong density was modelled as the inhibitor was observed adjacent to N2, helping confirm the position of the quinone cavity (Zickermann et al., 2015).

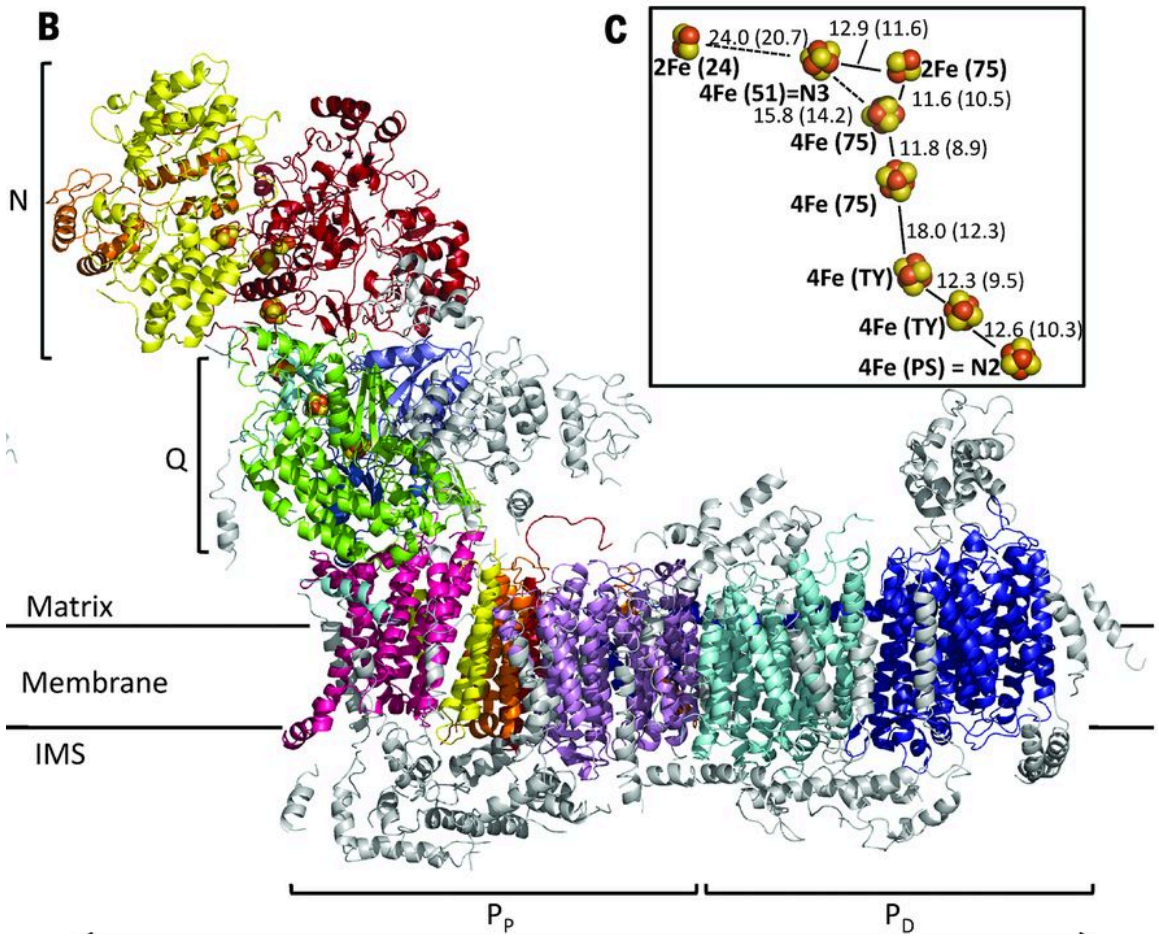


Figure 1.17: Structure of complex I from *Y. lipolytica*. (B) Complex I is displayed in cartoon representation, in which secondary structures care coloured according to their assigned subunits, while unassigned secondary structure elements are grey. The N (NADH oxidising domain), Q (ubiquinone domain), P_P (proximal membrane domain) and P_D (distal membrane domain) are indicated on the structure. (C) inset shows the iron-sulphur clusters with centre-to-centre (edge-to-edge) distances in Å. This figure was taken from Zickerman *et al.,* (Zickermann et al., 2015)

1.10 The structure of mammalian complex I

Solving the structure of the mammalian enzyme by X-ray crystallography has been notoriously difficult, with only the membrane domain (subcomplex Iβ) being resolved by this method (Zhu et al., 2015). The structure allowed for the assignment of subunits in the distal part of the membrane. However, considering the absence of most of the protein, little information could be gained about the organisation of the supernumerary subunits or the enzyme mechanism.

The first structure to establish the complete architecture of the eukaryotic complex I was that of the bovine enzyme by cryo-EM (Vinothkumar et al., 2014). At a global resolution of 4.95 Å, secondary structure features such as α helices and β sheets were apparent in the density, which led to the assignment of subunits into well-resolved densities (Figure 1.18). In addition to the 14 conserved core subunits, 14 supernumerary subunits were identified out of the 31 present (NDUFA1, NDUFA3, NDUFA5, NDUFA6, NDUFA8, NDUFA9, NDUFA10, NDUFA11, NDUFA13, NDUFAB1-α and NDUFAB1-β). The supernumerary subunits enclose most of the core subunits. The flavin domain, which is the site of superoxide production is exposed, therefore the function of the supernumerary subunits is unlikely to be the mitigation of superoxide production. An interesting feature is the presence of two copies of NDUFAB1 in the structure, one at the base of the matrix arm, and the other near the end of the membrane arm. This explains why NDUFAB1 was found in separate fractions in dissociation experiments (Carroll et al., 2003). The reason why there are two copies of NDUFAB1 on complex I is unclear at present. Also, amongst the supernumerary subunits modelled in is NDUFA9, which belongs to a family of short chain fatty acid dehydrogenases. NDUFA9 has a bound NADPH, confirmed by mutational studies on the yeast *Y. lipolytica* (Abdrakhmanova et al., 2006). The function of this nucleotide is unknown. Although abundant in structural information, the resolution of the structure (5 Å) precluded the assignment of amino acid residues, with exception of bulky aromatic residues. Therefore, most of the residues were modelled in as poly-alanines, which prevented analysis on key residues in mechanistically important sites.

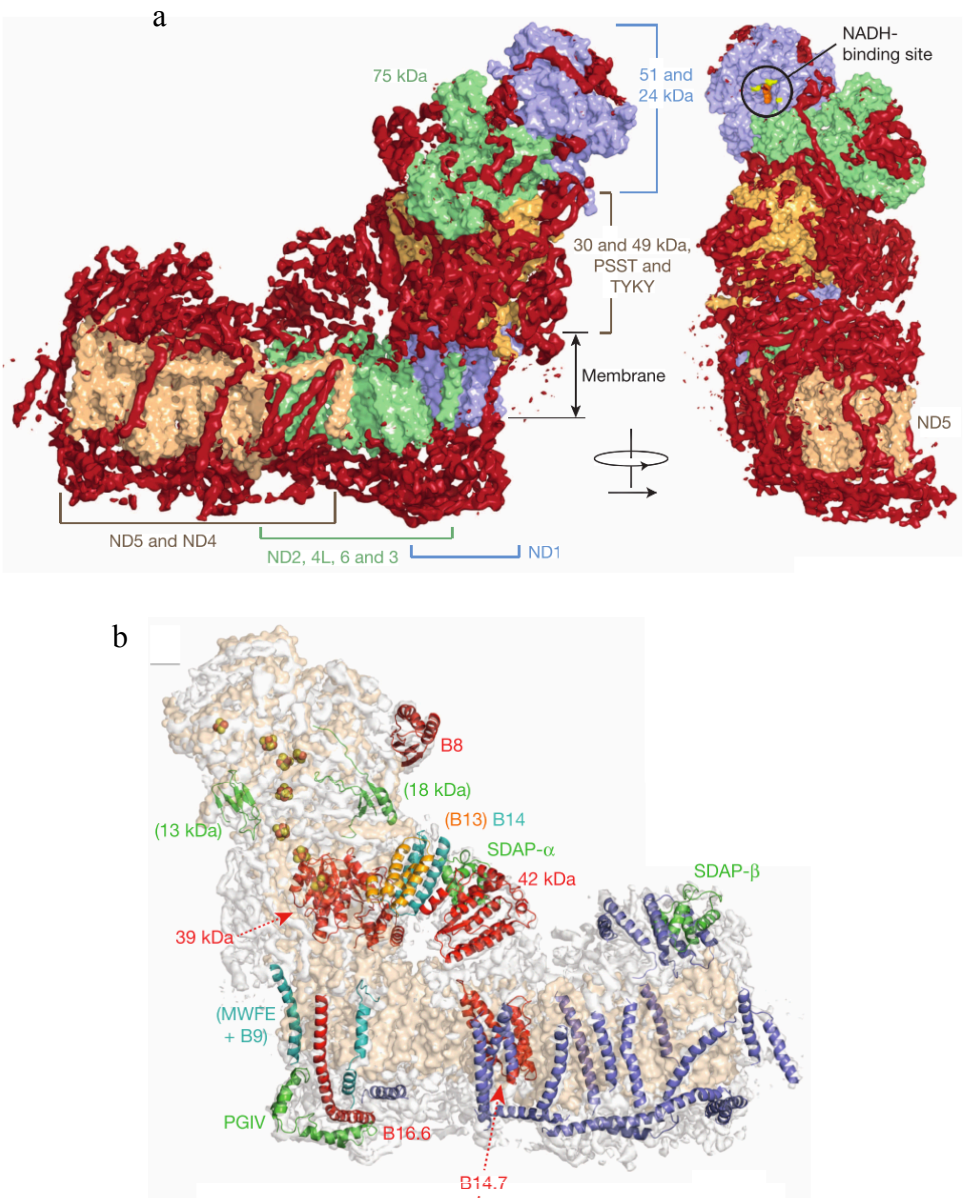


Figure 1.18: Architecture of bovine complex I. (a) The arrangement of supernumerary subunits (red) around the core subunits labelled (beige, green and blue). (b) A cartoon representation of assigned supernumerary subunits labelled, with core subunits and unassigned supernumerary densities shown in beige and grey, respectively. See table 1.1 for complex I subunit nomenclature. This figure was taken from Vinothkumar et al., (Vinothkumar et al., 2014).

The first complete structure of mammalian complex I was solved using cryo-EM on the bovine enzyme (Zhu et al., 2016b). The cryo-EM map was resolved globally to 4.16 Å, which allowed the modelling of all subunits into the structure (14 core and 31 supernumerary subunits) (Figure 1.19). The near-atomic resolution structure had many well-resolved features, such as side chain

densities, improving on the number of assigned residues from the previous bovine structure (Vinothkumar et al., 2014).

Interestingly, the supernumerary subunits in the membrane domain form a cage-like structure around the core subunits (with many disulphide bonds present in the subunits on the inter-membrane-space side), whilst the supernumerary subunits in the hydrophilic domain form threading structures, probably stabilising the protein by making large surface contacts with the core subunits. The contrast in the strategy for stabilisation of the core subunits by the two sets of supernumerary subunits may have some relevance in complex I assembly.

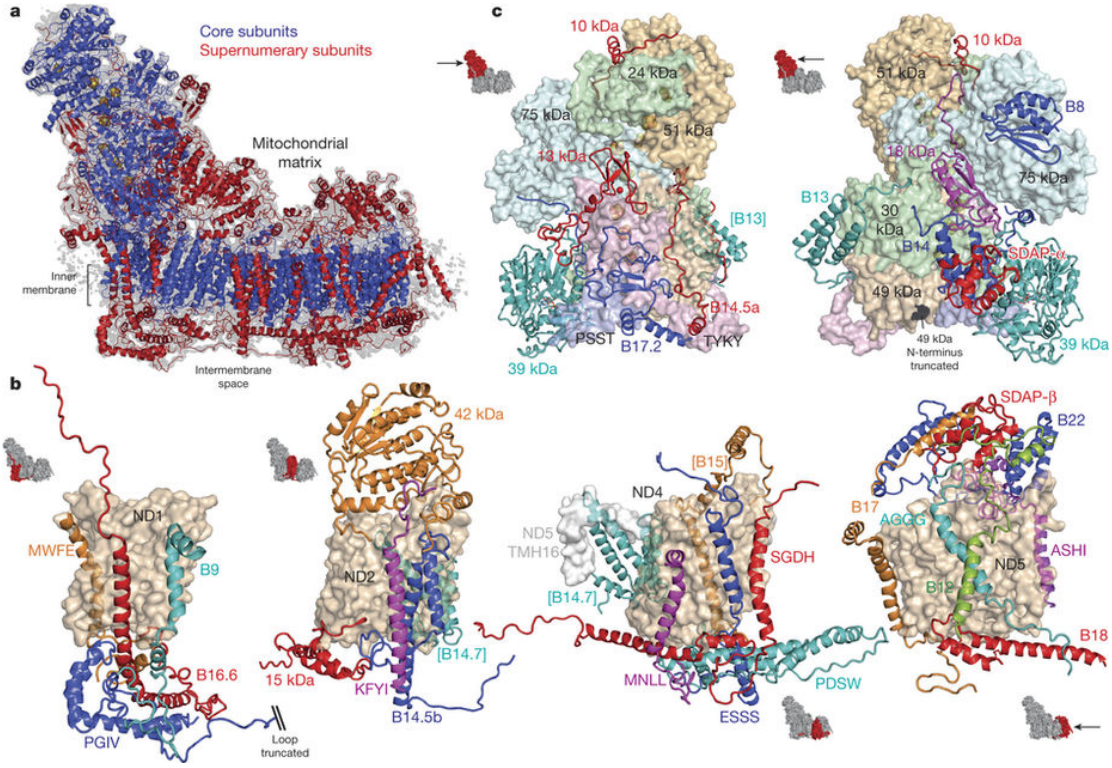


Figure 1.19: Structural overview of bovine complex I. (a) A cartoon representation of assigned supernumerary subunits (red) and core subunits (blue). (b and c) Positions of the supernumerary subunits (cartoon) in different domains of complex I (coloured red on the complex I icon), with core subunits presented as surface representations. See table 1.1 for complex I subunit nomenclature. This figure was taken from Zhu *et al.*, (Zhu et al., 2016).

Furthermore, 3D classification of the bovine particles resulted in three distinct classes, a minor broken class termed 'class 3', and two major classes termed 'class 2' and 'class 1'. The class 2 map (4.35 Å) and class 1 map (4.27 Å) revealed distinguishing features between the two, such

as disorder in the ubiquinone binding site and the ND3 loop between transmembrane helix 1 and 2 in the class 1 map. The regions of disorder in the class 1 map were unexpected considering it was of a higher resolution than the class 2 map. Class 1 was therefore proposed to be as the de-active state of complex I because of key critical features in the structure that were consistent with biochemical assays, such as an exposed cysteine 39 on the disordered ND3 loop, which is modified by N-ethylmaleimide (NEM) when complex I is de-active state (Galkin et al., 2008). In contrast, the ND3 loop and ubiquinone binding site in were ordered in class 2, including the ubiquinone binding site of complex I, therefore this class was proposed to be the active state of complex I. This is the first case in which biochemically-defined states were structurally described in complex I. This was later confirmed by solving the structure of de-activated bovine complex I (Blaza et al., 2018).

Shortly after the bovine complex I structure was solved, the *Ovis aries* (ovine) complex I structure was resolved to 3.9 Å (Fiedorczuk et al., 2016). In the ovine structure, all of the subunits were modelled into the structure with the aid of cross-linking mass spectrometry. The assignment and organisation of subunits were identical to that of the bovine enzyme. The ovine complex I also featured 12 bound phospholipids, which were previously shown to be critical for the activity of the enzyme (Sharpley et al., 2006) (Figure 1.20). The authors have suggested that the positioning of the lipids between hydrophobic subunits, in particular, cardiolipin, compensate for missing helices that are absent in metazoans (Birrell and Hirst, 2010), and therefore supposedly act to mitigate proton leak.

Initially, the ovine structure was classified as the de-active state of the enzyme because of the loop in NDUFS2 that is precluding ubiquinone access to the N2 cluster, thus preventing electron transfer. However, careful inspection of the ovine complex I density map (EMD-4084) reveals poor density for NDUFA11, indicative of partial occupancy for that subunit, and there is no apparent density for the second-half of the transverse helix in ND5 (Blaza et al., 2018). Moreover, comparison between the ovine map and known bovine complex I classes show that ovine complex I belongs to the broken state (class 3) of the enzyme. This is confirmed by the poor NADH:DQ activity displayed by the ovine enzyme (3.5 μmol min^{-1} mg^{-1}) compared to its bovine counterpart (24 μmol min^{-1} mg^{-1}) (Blaza et al., 2018).

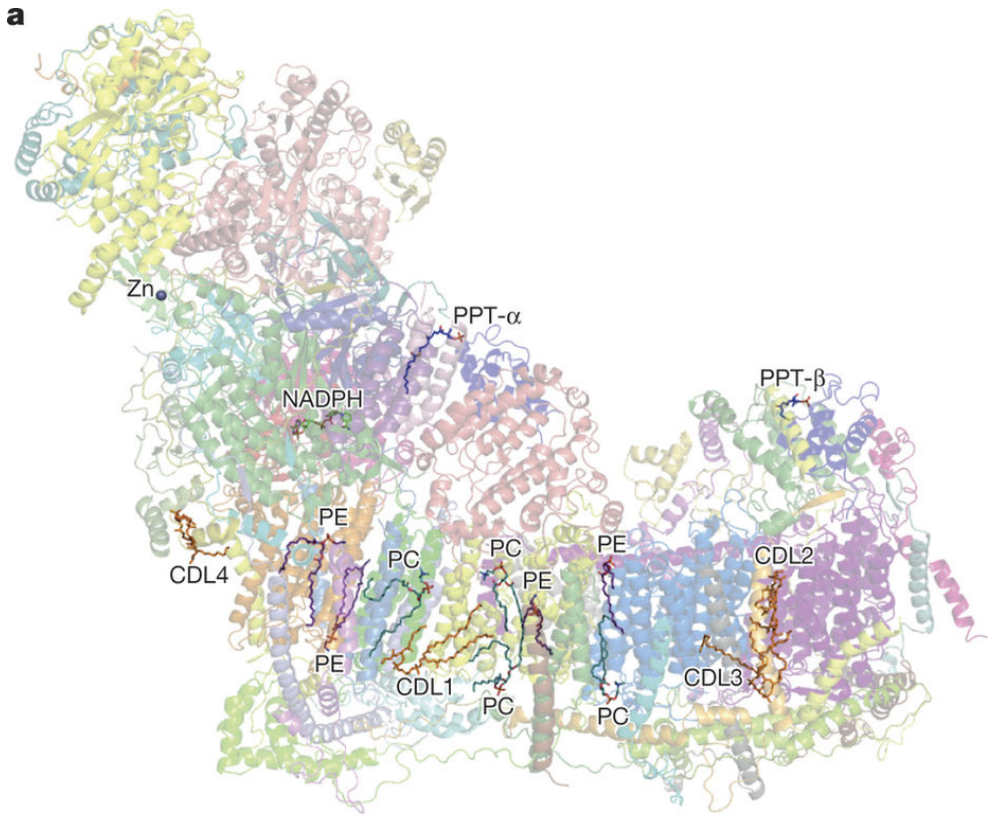


Figure 1.20: Positions of phospholipids in ovine complex I. Some co-factors are shown as sticks such as PC (phosphatidylcholine), PE (phosphatidylethanolamine), CDL (cardiolipin), NADPH, and PPT (phosphopantetheine), whilst Zn (zinc) is shown as a sphere. This figure was taken from Fiedorczuk *et al.,* (Fiedorczuk et al., 2016).

1.11 The structure of complex I in respiratory supercomplexes

The arrangement of respiratory complexes I, II and IV into supercomplexes in eukaryotes was first demonstrated by BN-PAGE analysis (Schägger and Pfeiffer, 2000). The authors also found that the vast majority of complex I in native membranes exists in supercomplexes rather than as isolated enzymes. In addition, there are three types of known supercomplexes in mammals, one without complex I (complexes III_2 and IV), and two with complex I (the respirasome which contains complexes I, III_2 and IV, and supercomplexes containing only complexes I and III_2) (Milenkovic et al., 2017).

The architecture of the respirasome was solved independently at the same time from both ovine and *Sus scrofa* (porcine) mitochondria at 5.8 Å and 5.4 Å, respectively (Gu et al., 2016; Letts et al., 2016a). The structure revealed that complex I interacts with complex III (specifically subunit UQCRQ) via NDUFA11 and NDUFB4, both found anchored to the transverse helix of ND5, and with complex IV (subunits COX7c and COX8b) via ND5 and NDUFB3, both at the tip of the membrane domain (Figure 1.21). In addition, two different conformations of the respirasome were found by Sazanov and co-workers, a 'tight' conformation where complex IV interacted more with complex I and III, and a 'loose' conformation where complex IV interacted less with complex I and III. Recently, there have been higher resolution structures of the porcine respirasome and the architecture of the human megacomplex has also been described (complexes I_2, III_2 and IV_2) (Guo et al., 2017; Wu et al., 2016a). However, the human megacomplex has not been previously observed by BN-PAGE.

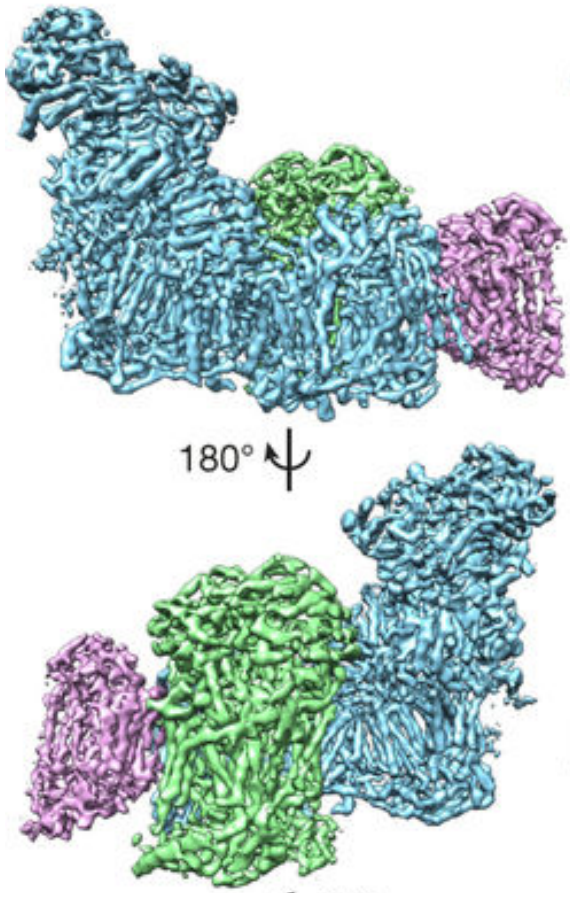


Figure 1.21: The arrangement of complexes in the respirasome. Cryo-EM density map of complex I is shown in blue, complex III dimers are shown in green and complex IV is shown in pink. The structure was taken from Letts *et al.,* (Letts et al., 2016a).

Respiratory supercomplexes have been proposed to be involved in substrate channelling of mobile electron carriers (cytochrome *c* and ubiquinol) between the complexes (Blanchi et al., 2004), and that the supercomplexes specifically have dedicated pools of electron carriers that are not exchanged with other enzymes outside the supercomplex (Lapuente-Brun et al., 2013). However, this is not supported by kinetic evidence (Blaza et al., 2014) and it is clear from the structure of the respirasome that the cytochrome *c* and ubiquinone/ubiquinol binding sites are accessible to the outside environment, suggesting the electron carriers are free to diffuse (Figure 1.22). Currently, it is not clear what is/are the function(s) of supercomplexes.

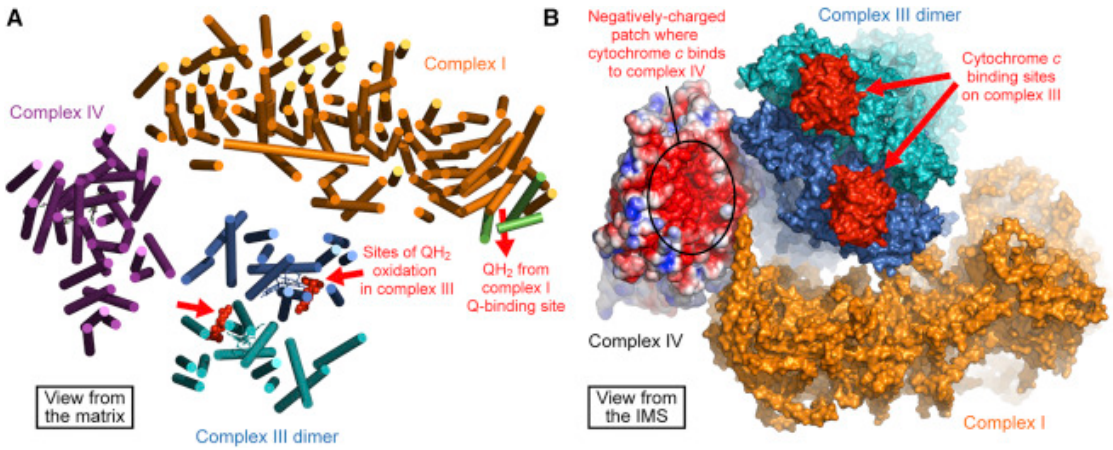


Figure 1.22: Ubiquinone and cytochrome c binding sites in the porcine respirasome. (a) The transmembrane helices are shown for individual complexes in the respirasome, with the ubiquinone/ubiquinol binding sites indicated. (b) The individual complexes are coloured differently with the binding site of cytochrome c indicated. This figure was taken from Milenkovic *et al.,* (Milenkovic et al., 2017).

1.12 Mammalian complex I Assembly

Mammalian complex I is a complicated protein requiring many assembly factors to mature. Assembly factors are only transiently associated with the nascent complex (Andrews et al., 2013; Heide et al., 2012). Complex I assembly has been proposed to occur by a sequential process, where individual modules, containing subunits from a particular domain are first assembled (Tuschen et al., 1990), and then these modules are joined together to form the mature complex I (Guerrero-Castillo et al., 2016).

Starting from the ubiquinone domain (Q-module), NDUFS2, NDUFS3 and NDUFA5 are the first subunits from the hydrophilic domain of complex I to assemble (Andrews et al., 2013; Mckenzie and Ryan, 2010). The association of these subunits with each other is consistent with their position in the structure. This subassembly is then joined by NDUFS7 and NDUFS8, along with assembly factors NDUFAF4, NDUFAF3 and TIMMDC1 (Andrews et al., 2013). The subassembly then anchors itself to ND1, following the addition of further subunits (NDUFA3, NDUFA8 and NDUFA19).

The second independent assembly module, the proximal membrane module (the membrane domain closer to the Q-domain) consists of subunits ND2, NDUFC1, NDUFC2 and assembly factors NDUFAF1, ECSIT, ACAD9, TMEM1126B and COA1 (Guerrero-Castillo et al., 2016). Mitochondrial encoded subunits ND6, ND4L and ND3 then join the proximal membrane domain.

The third independent assembly module, the distal module domain consists of mitochondrial encoded ND5 and NDUFAB1, NDUFB7, NDUFB3, NDUFB5, NDUFB9 and NDUFB2. The distal membrane module then joins the proximal membrane module, with the aid of assembly factors FOXRED1, ATP5LL and TMEM70 (Guerrero-Castillo et al., 2016). This then joins with the previously assembled Q-module.

Finally, the maturation of the complex is completed by the addition of the N-module (NDUFV1, NDUFV2, NDUFS1, NFUFA2, NUDFS4, NDUFV3, NDUFA12 and NDUFS6), and in the process, the assembly factors dissociate (see Figure 1.23 for the assembly scheme of mammalian complex I).

Although the stepwise assembly complex I has been characterised, the identification of the steps has relied only on BN-PAGE analyses and stable isotope labelling in cell cultures. Confirmation of the process will require future structural determination of the different assembly state intermediates. This may give insight into the function of these assembly factors.

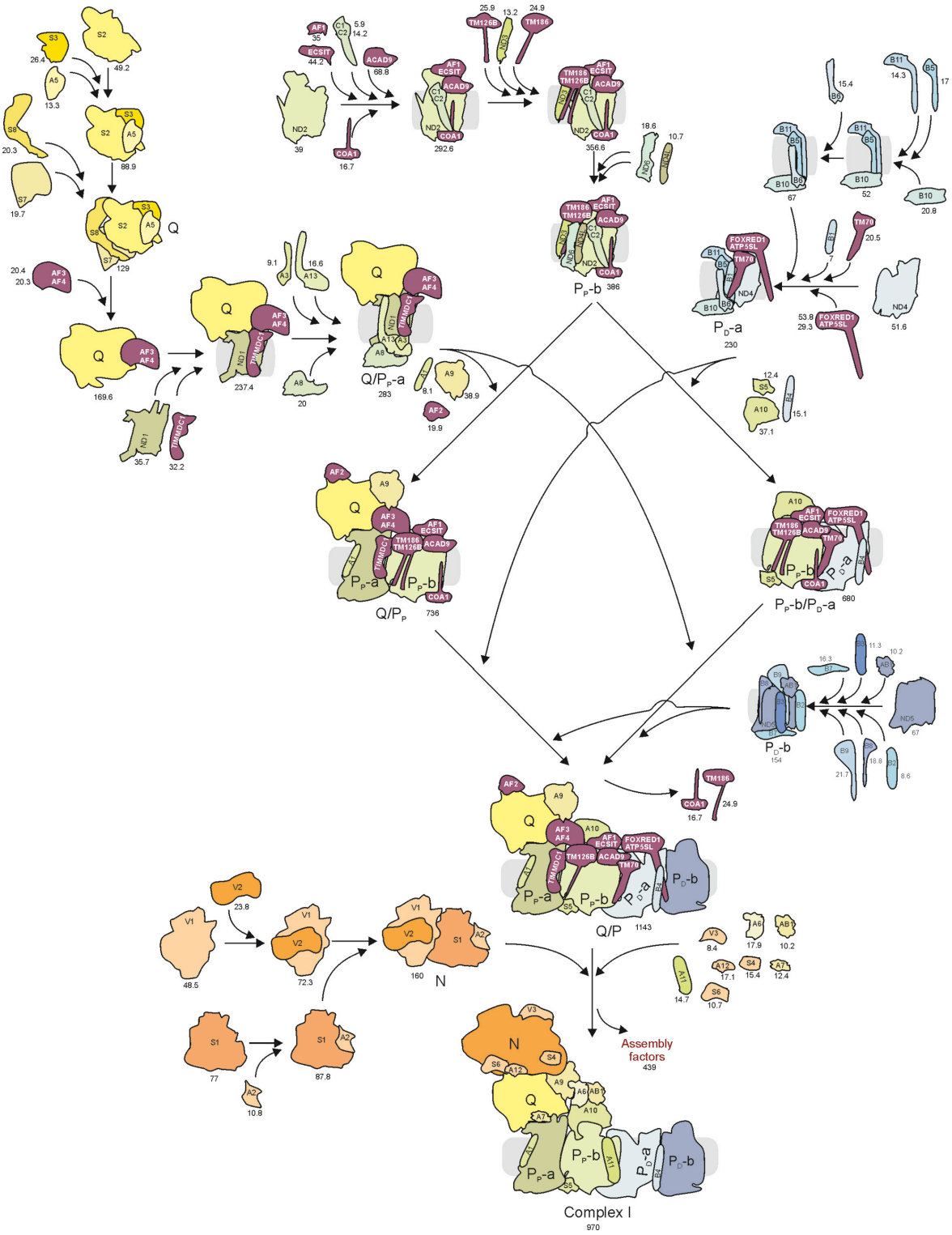


Figure 1.23: Stepwise assembly of complex I in human cell culture. Proteins from different domains of complex I are coloured differently, with extrinsic assembly factors coloured purple. This figure was taken from Guerrero-Castillo *et al*., (Guerrero-Castillo et al., 2016).

1.13 Complex I deficiencies

Complex I is encoded on both the nuclear genome and the mitochondrial genome, therefore mutations in either can result in complex I deficiencies in patients. In addition, mutations in the assembly factors of complex I can be detrimental to the maturation of the enzyme (Fassone et al., 2011) Complex I disorders are the most prevalent of mitochondrial disorders, accounting for ~30% of all known cases (Fassone and Rahman, 2012).

The clinical phenotype for complex I mutations vary greatly, making them difficult to diagnose; for example, the case of two siblings with the same mutation in *NDUFS4* but displaying different clinical features (Budde et al., 2003). In this example, the sister suffered from the typical neuromuscular phenotype displayed by patients with Leigh syndrome, while the brother had cardiomyopathy.

The most prevalent complex I disorder is Leigh syndrome, and it is an early-onset neuromuscular disorder, accounting for 39% of cases (Fassone and Rahman, 2012). Primarily, as with other complex I deficiencies, Leigh syndrome in patients is marked with the reduction in catalytic activity and the content of complex I in mitochondria. As a result, patients typically display symmetrical bilateral lesions in the basal ganglia, increased lactic acid in the blood plasma, ataxia, lethargy and ultimately death with the progression of the disease (Ortigoza-Escobar et al., 2016). An example, a homozygous mutation in NDUFS1 (L231V mutation) in an 8.5 month-old toddler resulted in bilateral lesions in the brain, ataxia, growth retardation and a reduced complex I activity (Martín et al., 2005).

Other less prevalent clinical manifestations of complex I deficiencies include, but not limited to, cardiomyopathy, leukoencephalopathy and fatal infantile lactic acidosis, with the latter causing death in infancy (Fassone and Rahman, 2012).

Unfortunately, due to the clinical heterogeneity displayed by patients with complex I deficiency and the overlap in phenotype with mutations in other mitochondrial proteins (Zhu et al., 1998), the only methods of diagnosis are by sequencing the genome(s) of patients to identify possible pathogenic mutations, or to perform biochemical and/or protein-content analyses from patient samples. In one such case, Mootha and co-workers sequenced a number of candidate genes from 103 patients with known complex I deficiency and compared the results to 42 healthy

individuals, thus identifying pathological mutations, such as those for complex I assembly factors *Nubpl* and *Foxred1* (Calvo et al., 2010). This was confirmed by rescuing complex I deficiency in patient fibroblast using cDNA carrying the 'healthy' genes (Calvo et al., 2010).

1.14 Aims for this thesis

Currently, model systems for complex I studies range from bacterial and yeast systems, which are genetically amenable but not medically-relevant, to large mammalian organisms such as *Bos taurus*, which are more closely related to humans but in which gene modification is not practical. In this thesis, *Mus musculus* (mouse) is developed as a system for complex I studies, because it is a mammalian, biomedically-relevant model system for human disease, which is amenable to genetic manipulation.

- The first part of this thesis describes the purification of complex I from mouse heart in a stable, active and pure state
- The second part of this thesis describes structural studies on the mouse enzyme using cryo-EM.

Mutations in complex I, in both the nuclear encoded and the mitochondrial encoded subunits lead to a plethora of neurological diseases, such as Leigh syndrome. The 'NDUFS4' mouse model, from which the NDUFS4 subunit of complex I has been deleted, recapitulates some of the effects seen in humans.

- The final part of this thesis aims to characterise the enzyme from the NDUFS4 disease model, and determine the effects of the deletion on the structure and activity of the complex.

Chapter 2: Materials and methods

2.1 Materials

All materials were purchased from Sigma-Aldrich or Thermo Fisher Scientific unless otherwise stated.

2.2 Genotyping the NDUFS4 knockout mouse strain

2.2.1 Genomic DNA extraction

DNA was extracted from *Mus musculus* (mouse) C57BL/6 strain ear biopsies using a 7 well Maxwell® 16 DNA Purification cartridge. To start automated genomic DNA extraction, the following selections were made sequentially on the Maxwell® 16 Instrument to start the program: Run, DNA and Tissue. After 45 min., extracted genomic DNA was transferred into 300 µL of nuclease-free H_2O.

2.2.2 Polymerase chain reaction (PCR)

PCR was performed according to Kruse *et al.*, (Kruse et al., 2008). In a total volume of 25 µL, the following additions were made: 5µL of 5X Green GoTaq® Reaction Buffer (Promega), 2 µL of 2.5 mM dNTPs, 1.25 µL dimethyl sulphoxide (DMSO), 0.6 µL of 20 µM forward primer, 0.6 µL of 20 µM reverse primer, 13.4 µL H_2O, 2 µL DNA and 0.15 µL GoTaq® G2 DNA Polymerase. The conditions and primers used for PCR are given below.

The primers used to amplify the undisrupted (wild-type) NDUFS4 gene. Expected product size: ~150 base pairs.
Forward primer: 5'- AGCCTGTTCTCATACCTCGG-3'
Reverse primer: 5'- GCTCTCTATGAGGGTACAGAG-3'
The primers used to amplify the modified (deletion of exon 2 resulting in the loss of the protein) NDUFS4 gene: Expected product size: ~200-300 base pairs.
Forward primer: 5'- AGCCTGTTCTCATACCTCGG-3'

Reverse primer: 5'- GGTGCATACTTATACTACTAGTAG-3'

PCR cycling parameter setup on the PCR instrument:

1. 95 °C for 3 min. (denaturing).
2. 95 °C for 30 sec. (denaturing).
3. 53 °C for 30 sec. (annealing).
4. 72 °C for 45 sec. (extension).
5. Repeat steps 2 - 4 an additional 34 times.
6. 72 °C for 5 min. (extension).
7. 4 °C indefinitely (storage).

2.2.3 Agarose gel electrophoresis

Gels for electrophoresis were prepared by dissolving 2% (w/v) agarose (BioGene) in Tris-Borate-EDTA buffer (TBE) by heating the mixture. 2 µL of ethidium bromide was added to the solution (50 mL) and left to solidify on a gel tray. 5 µL of the amplified DNA products were resolved at 100 V for 30 min. alongside a 1 kb plus DNA ladder (New England Biolabs). The fluorescence of ethidium bromide (which intercalates to DNA) was then imaged on a ChemiDoc imaging system (Bio-Rad).

2.3 Preparation of FumC and MaeB for complex II assay

2.3.1 Plasmids

Decarboxylating malic enzyme (MaeB) and fumarase C (FumC) from *E. coli* were expressed using plasmids conferring ampicillin resistance (kindly provided by Dr. Andrew Jones), pET32:*maeB* and pASK40:*fumC* respectively (Bologna et al., 2007; Weaver et al., 1995), in overexpressing *E. coli* BL21 (DE3) strain (New England Biolabs).

2.3.2 *E. coli* Growth conditions and protein overexpression

Preparation of MaeB and FumC was carried out as described in Jones *et al*., (Jones and Hirst, 2013). Bacterial cultures were grown in a flask by shaking overnight at 37 °C in 50 mL of lysogeny broth (LB) media supplemented with 50 μg mL^{-1} of ampicillin. Then, ~1 mL of the starter culture was inoculated into 1 L of the same medium, and grown by shaking (150 rpm) at 32 °C. The optical density was monitored at 600 nm until an absorbance of 0.6 was reached. Protein expression was then induced by addition of 1 mM and 0.1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) for FumC and MaeB, respectively. The induced bacterial cultures were left to grow overnight at 20 °C, ready for cell harvesting.

2.3.3 Cell harvesting and protein purification

All procedures were carried out at 4 °C. To pellet the cells, bacteria were centrifuged at 7,000 × *g* for 15 min. in an RC12-BP centrifuge. Pelleted cells from 4 L cultures were then re-suspended in ~120 ml of buffer A1 (100 mM NaCl, 25 mM imidazole, 10% glycerol (v/v, VWR) and 20 mM Tris-SO_4 pH 7.4 at 4 ˚C) and buffer A2 (300 mM NaCl and 50 mM potassium phosphate pH 7.8 at 4 °C) for MaeB and FumC, respectively. Three EDTA-free protease inhibitor cocktail tablets (Roche) were added to impede proteolysis. Cells were then ruptured using a Z-Plus 2.2 kW cell disruptor (Constant Systems) at 30 kpsi, by passing them twice through the system. The lysate was then centrifuged at 160,000 × *g* for 45 min. using a Type 45 Ti rotor (Beckman Coulter). The resulting pellet was discarded and the supernatant was filtered with a 0.22 μM syringe filter (Millipore). Using a peristaltic pump, the filtered supernatant was loaded onto a 2 x 5 mL HisTrap HP column (GE Healthcare) at 2 mL $min.^{-1}$. The column was then attached to an Äkta prime FPLC system (GE Healthcare), and washed with buffer A1 and A2 for FumC and MaeB respectively, until the absorbance at 280 nm The imidazole concentrations were increased over 40 mL with buffer B1 (buffer A1 + 300 mM imidazole) and buffer B2 (buffer A2 + 400 mM imidazole) for MaeB and FumC, respectively. The histidine-tagged proteins both eluted in ~75-80% imidazole of the eluent buffers. The fractions were pooled and concentrated 10-fold using 50 kDa molecular weight cut-off (MWCO) centrifugal filter concentrators (Millipore). The concentrated proteins were then dialysed for 12 hr. in buffer B1 (20 mM β-mercaptoethanol, 10% glycerol and 60 mM Tris–SO_4 pH 7.4 at 4 ˚C) and buffer B2 (5 mM EDTA (VWR) 1 mM 1,4-Dithiothreitol (DTT) and 10 mM tris–SO_4 pH 7.0 at 4 ˚C) for MaeB and FumC respectively. Finally, the proteins were

stored at -80 ˚C in 50% glycerol. The yields obtained from 4 L of bacterial culture were 144 mg and 54 mg of protein for MaeB and FumC, respectively.

2.4 Preparation of mitochondrial membranes

2.41 Heart extraction

574 C57BL/6 mice were sacrificed by cervical dislocation, in accord with the UK Animals (Scientific Procedures) Act, 1986 (PPL: 7538, approved by the local ethics committee and the UK Home Office). Hearts (typically 0.2 g per heart) were surgically removed, then immersed in cold AT buffer (10 mM Tris-HCl pH 7.4 at 4 °C, 75 mM sucrose, 225 mM sorbitol 1 mM EGTA and 0.1% (w/v) fatty acid-free bovine serum albumin).

2.4.1 Preparation of mitochondria from hearts

Mitochondria were isolated from hearts using methods based on previous protocols (Fernández-vizarra et al., 2010; Fernández-Vizarra et al., 2002). In order to ease the homogenisation process, hearts were cut into 1 mm, then washed several times with AT buffer to remove blood and other hair. Heart pieces were then re-suspended in AT buffer (10 mL per g of material) and homogenised to consistency using a Potter-Elvehjem homogenizer fitted with a Teflon pestle, by 7 - 10 strokes at ~1,000 rpm (IKA® RW20 Digital overhead stirrer), in the case of mouse hearts. The homogenate was centrifuged at 1,000 x *g* for 5 min. to separate the mitochondria from cellular debris. The supernatant was carefully decanted and centrifuged at 9,000 x *g* for 10 min. to pellet crude mitochondria. Mitochondria were re-suspended with AT buffer 1/8th of the starting volume and re-centrifuged at ~15,000 x *g* for 2 min. This process was repeated several times to remove fatty acids. Finally, pelleted mitochondria were re-suspended in mitochondrial re-suspension buffer (20 mM Tris-HCl pH 7.4 at 4 °C, 1 mM EDTA, and 10% glycerol) to 10-20 mg mL^{-1} of protein and stored at -80 °C.

2.4.2 Isolation of mitochondrial membranes

Mitochondria were thawed and diluted to 5 mg mL^{-1} with mitochondrial re-suspension buffer, then ruptured by three 5 sec. bursts of sonication (with 30 sec. intervals on ice) using a Q700 Sonicator (Qsonica) at 65% amplitude setting. The membranes were pelleted at 75,000 x *g* using an MLA80 rotor (Beckman Coulter) for 1 hr., then re-suspended in mitochondrial re-suspension buffer to 2/3rd of the starting volume to give ~5 mg mL^{-1} of protein.

2.4.3 Origin of *Ndufs4* ($^{-/-}$) mice

Ndufs4 ($^{-/-}$) mice were generated by Palmiter and co-workers (Kruse et al., 2008). In brief, the second exon of the *Ndufs4* gene was flanked by loxP sites in AK18.1 mouse embryonic stem cells (ESC). The resulting genetically modified mice were then crossed with Mox2-Cre to remove the second exon from the *Ndufs4* gene. Mice that were heterozygous for truncation in the *Ndufs4* allele were interbred to produce *Ndufs4* ($^{-/-}$) mice.

2.5 Protein visualisation techniques

2.5.1 BN-PAGE

The following method was carried out at 4 °C. Heart mitochondrial membranes (5 mg mL^{-1}) were solubilised by the addition of n-dodecyl β-D-maltoside (DDM, Glycon) or lauryl maltose neopentyl glycol (LMNG, Anatrace), to give a final detergent concentration of 1% (w/v). Following a 30 min. incubation, the samples were centrifuged at ~13,000 × *g* for 30 min. using a benchtop microfuge. The resulting supernatant (solubilised membranes) was analysed by blue native PAGE (BN-PAGE) according to previous methods (Schägger and von Jagow, 1991; Wittig et al., 2006). In brief, solubilised membranes were diluted by 1/4th with 4 x NativePage^{TM} sample buffer (Invitrogen), and then Commassie G-250 (Amnesco) was added to give a dye/detergent ratio of 1/4. The samples were then loaded onto NativePage^{TM} 3 - 12% Bis-Tris gels (Invitrogen). Initially, electrophoresis was carried out at 80 V for 30 min. using NativePage^{TM} (Invitrogen) inner chamber buffer (1 x cathode buffer additive and 1 x running buffer) and outer chamber buffer (1 x running buffer). After 30 min, the inner chamber buffer was exchanged with 0.1 x cathode buffer additive and 1 x running buffer, and then electrophoresis was carried out at 180 V for an additional 2 hr. In-gel NADH oxidation was visualised by incubating native gels with 0.5 mg mL^{-1} Nitroblue tetrazolium chloride (NBT)

and 100 µM NADH in 20 mM Tris-HCl pH 7.55 at 32 °C for 20 min. NADH oxidation is denoted by the appearance of purple bands. Native gels not used for in-gel activity detection were stained with 50% methanol (v/v), 7% acetic acid (v/v) and 0.2% Commassie R250 (w/v) for 5 min, then de-stained overnight with 20% methanol and 7% acetic acid.

2.5.2 SDS-PAGE

Samples for SDS-PAGE were incubated in 100 mM DTT, 63 mM Tris-HCl pH 6.8 at 25 °C, 2% Sodium dodecyl sulphate (w/v, SDS), 10% glycerol and 15 nM bromophenol blue. The samples were loaded onto a Novex™ 10 - 20% Tris-Glycine gel (Invitrogen) and subjected to electrophoresis at 125 V for 90 min. Gels not used for Western blotting were stained and de-stained similarly as the native gels.

2.5.3 Western blotting and immunodetection

SDS-PAGE gels were transferred to Immobilon®-FL polyvinylidene fluoride membrane (Millipore) using a wet transfer system in 25 mM Tris, 190 mM glycine and 10% methanol at 300 mA for 1 hr. The membrane was then blocked for 1.5 hr. with the Odyssey (Li-COR Bioscience) buffer, then stained overnight at 4 °C with either polyclonal (anti-rabbit) NDUFB9 antibody (Proteintech: 15572-1-AP, 2000-fold dilution in Odyssey buffer) or polyclonal (anti-rabbit) NDUFS1 antibody (in-house, 10000-fold dilution in Odyssey buffer) that was kindly provided by Dr. Ian Watt from the MRC-MBU, University of Cambridge. The membranes were then incubated with a fluorescent secondary anti-rabbit antibody (IRDye® 680RD, Li-COR Bioscience) and imaged using a Li-COR imaging system (emission at 694 nm).

2.6 Kinetic activity assays

All kinetics measurements were conducted using a 96-well Spectramax 384 plate reader at 32 °C. Unless stated otherwise, all reactions were carried out in 20 mM Tris-Cl pH 7.55 at 32 °C as described previously (Birrell et al., 2009). An overview of the kinetic assays carried out is presented in Table 2.1.

Additions	HAR	FeCN	APAD	ATP hydrolysis	NADH:O_2	Succinate:O_2	H_2O_2 C1	C2	C3	C4
Alamethicin (µg mL^{-1})	15*	15*	15*	15	10	10	15*	15*	15*	15*
Piericidin (µM)	2*	2*	2*	-	-	-	2*	2*	2*	2*
Protein (µg mL^{-1})	1* or 0.1†	1* or 0.1†	10* or 1†	5‡	10	10	20* or 8†	20* or 8†	-	-
NADH(µM)	100‡	100‡	100‡	200	100‡	-	30‡	30‡	30‡	30‡
NADP$^+$ (mM)	-	-	-	-	-	2 ‡	-	-	-	-
Cytochrome c (µM)	-	-	-	-	1.5	1.5	-	-	-	-
HRP (U mL^{-1})	-	-	-	-	-	-	10	10	10	10
SOD (U mL^{-1})	-	-	-	-	-	-	10	10	10	10
Amplex Red (µM)	-	-	-	-	-	-	10	10	10	10
Catalase (U mL^{-1})	-	-	-	-	-	-	-	5000	-	5000
ATP (µM)	-	-	-	200	-	-	-	-	-	-
PEP (µM)	-	-	-	400	-	-	-	-	-	-
Lactate dehydrogenase (µg mL^{-1})	-	-	-	40	-	-	-	-	-	-
Pyruvate kinase (µg mL^{-1})	-	-	-	50	-	-	-	-	-	-
$MgSO_4$ (mM)	-	-	-	2	-	-	-	-	-	-
K_2SO_4 (mM)	-	-	-	2	-	-	-	-	-	-

Table 2.1: Overview of kinetic assays. * and † denote additions when membranes or complex I are used in the assays respectively. ‡ denotes the initiator of the reaction. C1-C4 represent the 4 conditions in the H_2O_2 production assay. All reactions were monitored at 340 – 380 nm (ε = 4.81 mM^{-1} cm^{-1}), except APAD and H_2O_2, which were monitored at 400 – 450 nm (ε = 3.16 mM^{-1} cm^{-1}) and 557 – 620 nm (ε = 51.60 mM^{-1} cm^{-1}), respectively.

2.6.1 Flavin-site NADH oxidation assays

Flavin-site NADH oxidation reactions were initiated by the addition of 100 µM NADH to the assay. Nucleotide-bound dependent HAR (hexaammineruthenium (III) chloride) reduction activity by membranes or purified complex I was measured at a final concentration of 1 µg mL^{-1} or 0.1 µg mL^{-1} respectively, using 3.5 mM HAR (Santa Cruz Biotechnology). NADH oxidation was monitored at 340 – 380 nm (ε = 4.81 mM^{-1} cm^{-1}). Potassium hexacyanoferrate (III) (FeCN) reduction activity by membranes (1 µg mL^{-1}) or complex I (0.1 µg mL^{-1}) using 1 mM FeCN was monitored similarly by following the NADH absorbance. 3-Acetylpyridine adenine dinucleotide ($APAD^+$) reduction activity by membranes (10 µg mL^{-1}) or complex I (1 µg mL^{-1}) using 0.5 mM $APAD^+$ was monitored by the reduction of $APAD^+$ to APADH at 400 – 450 nm (ε = 3.16 mM^{-1} cm^{-1}).

2.6.2 H_2O_2 production

Hydrogen peroxide production was measured using the Amplex Red method according to Kussmaul *et al*., (Kussmaul and Hirst, 2006). To establish protein-specific superoxide production, four conditions were set up, all with 10 U mL^{-1} horseradish peroxidase (HRP), 10 µM Amplex Red and 10 U mL^{-1} superoxide dismutase from bovine erythrocytes (SOD).

Condition 1: membranes (20 µg mL^{-1}) or complex I (8 µg mL^{-1})
Condition 2: membranes (20 µg mL^{-1}) or complex I (8 µg mL^{-1}) and 5000 U mL^{-1} catalase from *Corynebacterium glutamicum*.
Condition 3: No additions.
Condition 4: Only 5000 U mL^{-1} catalase.

The reaction was initiated with 30 µM of NADH, and monitored by the oxidation of Amplex Red to resorufin at 557 – 620 nm (ε = 51.60 mM^{-1} cm^{-1}). The catalase-sensitive superoxide production rate of the protein was calculated using the following equation derived from previous work (Kussmaul and Hirst, 2006): catalase sensitive rate = condition 1 – (condition 2 + condition 3 – condition 4).

2.6.3 NADH and succinate:O_2 oxidoreduction assays

Membranes (10 µg mL^{-1}) were permeabilised with 10 µg mL^{-1} alamethicin and supplemented with 1.5 µM cytochrome *c* from horse heart. Activity measurements were carried out in buffer K (10 mM Tris-SO_4 pH 7.0 at 32 °C and 250 mM sucrose). NADH:O_2 (catalysis by complexes I, III and IV) oxidoreduction was initiated with 100 µM NADH and its oxidation monitored at 340 – 380 nm (ε = 4.81 mM^{-1} cm^{-1}). Succinate:O_2 (catalysis by complexes II, III and IV) oxidoreduction was carried out according to the spectrally-sensitive and stoichiometric method of Jones and Hirst (Jones and Hirst, 2013). In brief, membranes were assayed with 2 mM $MgSO_4$, 2 mM K_2SO_4, 5 mM succinate (pH ~7), 90 µg mL^{-1} FumC and 500 µg mL^{-1} MaeB. The reaction was initiated with 2 mM $NADP^+$ and its reduction monitored at 340 – 380 nm (ε = 4.81 mM^{-1} cm^{-1}). The reaction steps (and the enzyme in parenthesis) for the succinate:O_2 are given below:

1. Succinate and ubiquinone are converted to fumarate and ubiquinol (complex II).
2. Fumarate and H_2O are converted to malate (FumC).
3. Malate and $NADP^+$ are converted to pyruvate, CO_2 and NADPH (MaeB).

2.6.4 ATP hydrolysis

ATP hydrolysis was measured according to the method of Pullman *et al*., (Pullman et al., 1960). Membranes (5 µg mL^{-1}) and 15 µg mL^{-1} alamethicin added to a pre-incubated (~2min.) mixture of 200 µM ATP, 400 µM phosphoenolpyruvate (PEP), 50 µg mL^{-1} pyruvate kinase from rabbit, 200 µM NADH, 40 µg mL^{-1} lactate dehydrogenase from bovine heart, 2 mM $MgSO_4$ and 2 mM K_2SO_4. ATP hydrolysis was followed by the oxidation of NADH.

2.7 Purification of complex I

2.7.1 Solubilisation of mitochondrial membranes

All purification steps were carried out at 4 °C unless stated otherwise. The complex I purification method was adapted from that for bovine (Sharpley et al., 2006). Thawed membranes (5 mg mL^{-1}) were incubated with 0.005% PMSF (w/v) to inhibit proteolysis. DDM

(1%) was added to mouse membranes, then stirred on a Thermomixer Comfort (Eppendorf) at 400 - 600 RPM for 30 min. to solubilize the complexes. In the case of deactivated mouse complex I, the membranes (17 mg mL^{-1}) were pre-treated at 37 °C for 30 min., then diluted to 5 mg mL^{-1} with buffer C. To clarify the solution, the sample was centrifuged at 48,000 x *g* using an MLA80 for 30 min.

2.7.2 Anion-exchange chromatography

Solubilised membranes were loaded onto a 1 mL Hi-Trap Q HP anion exchange column (GE Healthcare), pre-equilibrated with buffer A (20 mM Tris-HCl pH 7.14 at 23 °C, 1 mM EDTA, 0.1% DDM, 10% ethylene glycol (v/v, VWR), 0.005% asolectin (Avanti) and 0.005% CHAPS (Calbiochem). The column was attached to an Äkta micro system (GE Healthcare), then washed with buffer A at 0.3 mL $min.^{-1}$ for 2 min. with absorbance measured at 280 nm and 420 nm. Initially, the system was washed with 20% buffer B (buffer A + 1M NaCl) for 7 min. to remove unwanted proteins, then with 35% buffer B for 10 min. to elute complex I-containing fractions, and finally with 100% buffer B for 5 min. to remove the remaining bound proteins. Complex I-containing fractions (~1 mL) were pooled and concentrated (~100 µL) using a 100 kDa MWCO Vivaspin 500 concentrator (Sartorius).

2.7.3 Size-exclusion chromatography

Complex I was further resolved by size-exclusion chromatography, with the absorbance measured at 280 nm and 420 nm. Pooled and concentrated Q-sepharose fractions (100 µL) were injected onto either a Superose 6™ Increase 3.2/300 column used in the earlier purifications, or a Superose 6™ Increase 5/150 (GE Healthcare) column used currently, both pre-equilibrated in buffer C (20 mM Tris-HCl pH 7.14 at 23 °C, 10% ethylene glycol, 150 mM NaCl and 0.02% DDM). Complex I elutes as a pure monodisperse peak at ~1.39 mL and ~1.65 mL from the 3.2/300 and 5/150 Superose 6™ Increase columns, respectively. Protein for cryo-EM grid preparation was not concentrated before use, but for other uses complex I was concentrated and stored in 30% glycerol at -80 °C. Note for cryo-EM preparations; ethylene glycol was omitted in buffer C, due to a predicted decrease in image contrast, affecting 3D reconstructions of complex I. In addition, the detergent concentration in buffer C was increased from 0.02 to 0.05% for membranes starting material with 10 mg or more of protein, to maintain reasonable detergent to protein ratio, thus avoiding protein aggregation.

2.8 Cross-linking experiments

For cross-linking experiments, 0.1 - 1 mM bis(sulfosuccinimidyl)suberate (BS^3) were added to membranes in mitochondrial re-suspension buffer 5 min. after DDM addition and then incubated for 25 min.

2.9 Mass spectrometry

Dr. Michael Harbour and Dr. Shujing Ding from the MRC-MBU, University of Cambridge carried out the mass spectrometry experiments.

2.9.1 Data processing

Peptide spectra were recorded in 'raw files' by the instrument's operating software, Xcalibur. Proteome Discoverer was then used to open the raw files, collect MS/MS fragmentation spectra, and submit them to Mascot 2.4 (Matrix Science Ltd.) for comparison against a given database, to assign spectra to peptide sequences and proteins. Only hits with a threshold $p < 0.05$ (representing the 95% confidence threshold) were considered. Proteome Discoverer, then assigned confidence values for detected peptides, green=high confidence, yellow=medium confidence and red=low confidence. The protein database used for protein identification in mouse was UP000000589 (January 2016), with 53,106 protein entries from the UniProt server (Apweiler et al., 2017). Possible protein isoforms, predicted using Augustus (Stanke and Morgenstern, 2005), were manually inspected and added to the databases.

2.9.2 Tryptic digestion

Gel slices were washed thoroughly with 100 µL H_2O, then placed into 20 mM Tris-HCl pH 8.0. The samples were incubated for 1 hr. in 20 mM Tris-HCl pH 8.0 and 50% acetonitrile (v/v, Fluka), then in 100% acetonitrile to desiccate the gel for 15 min. To remove all of the liquid, to assist absorption of trypsin, the gel was vacuum dried (Savant SPD121P SpeedVac concentrator). The proteins in the gel were then digested overnight at 37 °C with 10 µL of bovine trypsin (12.5 µg mL^{-1}) in 20 mM Tris-HCl pH 8.0 and 5 mM $CaCl_2$. The digested tryptic peptides were extracted from the gel into the tube for 1 hr. using 15 µL of 4% formic acid (v/v)

and 60% acetonitrile. The extracted solution was transferred into a new tube, and the gel was then re-extracted for 1 hr. and added to the previously extraction solution. In the case of native gels, the extracted solution was further vacuum dried to increase the concentration of peptides and remove excess acetonitrile, which would prevent peptides from adhering to the reverse-phase high-performance liquid chromatography column. Finally, it was re-suspended in 15 µL of 2% acetonitrile and 0.1% formic acid.

Protein solution (~200 ng) for digestion were precipitated overnight in 100 µL cold ethanol solution at -20 °C. Precipitates were pelleted at 16,000 x *g* for 5 min. using a microfuge, and then re-suspended in 10 µL of 50 mM ammonium bicarbonate. Thereafter, they were digested overnight in 10 µL of trypsin (12.5 µg mL^{-1}) in 20 mM Tris-HCl pH 8.0 and 5 mM $CaCl_2$.

2.9.3 Matrix-assisted laser desorption/ionization time-of-flight/time-of-flight (MALDI TOF/TOF)

The tryptically-digested solution (0.5 µL) was spotted onto an Opti-TOF 384 MALDI plate, and to that 0.5 µL of 10 µg mL^{-1} α-cyano-4-hydroxycinnamic acid (CHCA) in 60% acetonitrile and 0.1% trifluoroacetic acid (v/v) was added. Once the solution had crystallised, the samples (each spot) were analysed sequentially by a 4800 MALDI TOF/TOF™ Analyzer (Applied Biosystems/MDS SCIEX). The autolysis products of trypsin and the characteristic signal peak of the CHCA matrix were used to calibrate each spectrum. The 7 highest scoring peptides from each spot (between 900-4000 m/z, where m/z is the mass to charge ratio) were automatically selected by the instrument's (4000 Series Explorer) software, for collision-based fragmentation (with atmospheric air) or collision-induced-dissociation (CID). Fragmentation spectra were then used for probability-based identification with the Mascot search engine (see above). The Mascot search parameters used against the database were peptide mass tolerance=70 ppm, fragment mass tolerance=0.8 Da, variable protein modifications=protein N-terminal acetylation, protein N-terminal formylation, methionine oxidation and cysteine propionamide derivatisation, with a maximum of 1 missed cleavage.

2.9.4 Liquid chromatography–mass spectrometry (LC-MS/MS)

Peptides from in-solution digestion were resolved by reverse-phase high-performance liquid chromatography (RP-HPLC) on an Acclaim PepMap nanoViper C18 reverse-phase column (75 µM x 100 mm). With a flow rate of 300 nL $min.^{-1}$, buffer A (5% acetonitrile and 0.1% formic acid) and buffer B (95% acetonitrile and 0.1% formic acid) were used to set up a gradient up to 40% acetonitrile over 84 min. Peptides were introduced directly into the linear trap Orbitrap XL™ mass spectrometer by electrospray ionisation. The 10 highest scoring peptides detected (between 400-6000 m/z) in the survey scan were selected for collision-based fragmentation. To avoid repeated fragmentation of a particular peptide, a dynamic exclusion list was made for a specified time, to exclude peptides that have already been fragmented from further analyses. The Mascot search parameters used against the database were the following: peptide mass tolerance=5 ppm, fragment mass tolerance=0.5 Da, variable protein modifications=protein N-terminal acetylation, protein N-terminal formylation and methionine oxidation, with a maximum of 4 missed cleavages. The number of missed cleavages was raised from 1 to 4, to increase the number of peptides detected despite also increasing the number of false positives.

2.9.5 Molecular mass measurements

Molecular mass measurements of complex I subunits were performed according to previous work on the bovine enzyme (Carroll et al., 2009). Protein mixtures were precipitated overnight with cold ethanol solution (20 vol.) at -20 °C. The protein was collected by centrifugation at 16,000 x *g* for 5 min., then re-suspended in 30 µL of formic acid/trifluoroethanol/hexafluoro-propan-2-ol/H_2O at a volume/volume (v/v) ratio of 60:15:1:24, with 1 mM tris(2-carboxyethyl)phosphine hydrochloride. The sample was centrifuged again at 16,000 x *g* for 5 min and the supernatant separated on a polymeric reverse-phase column (50 mm x 1 mm i.d., 50 °C) using an Agilent 1100 series HPLC system. Proteins were eluted with a gradient of propan-2-ol at a flow rate of 50 µL $min.^{-1}$, using the following buffer system; buffer A (trifluoroethanol/hexafluoro-propan-2-ol/1 M ammonium formate pH 3.1/H_2O at a volume/volume (v/v) ratio of 15:1:5:79) and buffer B (trifluoroethanol/ hexafluoro-propan-2-ol/propan-2-ol/1 M ammonium formate pH 3.1/H_2O at a volume/volume (v/v) ratio of 20:1:70:5:4) were used to set up a gradient up to 100% buffer B over 5 mL. The eluted proteins were introduced by electrospray ionisation into a 4000 Q trap mass spectrometer. The molecular masses of subunits were calculated from their multiple charged species using

PeakView software (SCIEX). Assignment of the measured masses to the mature proteins were made with the aid of previous data on the bovine enzyme (Carroll et al., 2009).

2.9.6 Complexome profiling experiments

Complexome analyses was carried out according to Heide *et al*., (Heide et al., 2012), by excising 64 x 1 mm slices from BN-PAGE gels. BN-PAGE was performed as described in Chapter 2.5.1, and tryptic digestion carried out as described in Chapter 2.8.2. Tryptically-digested peptides were resolved using reverse-phase HPLC and directly introduced into a LTQ Orbitrap XL^{TM} mass spectrometer. The 10 highest scoring peptides (between 400 - 2000 m/z) in the survey scan were chosen for collision-based fragmentation. To maximise the range of peptides detected, a dynamic exclusion list was included to prevent repeated fragmentation of the same peptide within a specified time. The Mascot search parameters used against the database were the following: peptide mass tolerance = 10 ppm, fragment mass tolerance = 0.5 Da, variable protein modifications = protein N-terminal acetylation, protein N-terminal formylation, methionine oxidation and cysteine propionamide derivatisation, with a maximum of 1 missed cleavage. Proteome Discoverer was used to submit the fragmentation spectra to Mascot for the identification of the proteins in each gel slice, and then retrieve the results for each gel slice, to be compiled into a consensus report. The proteins from each gel slice were quantified (relative) based on averaging the intensity values of its 3 highest detected peptides (by the intensity of the peptide precursor peaks in the survey scan). Pearson uncentered correlation coefficient (PUCC) is then calculated for every possible pair of protein detected. The resulting PUCC values are used to cluster proteins according to the single-linkage hierarchical cluster method, where initially each protein is a cluster of its own and is thereafter added to other protein clusters if the coefficient values are similar. Once a protein is clustered, the averaged relative protein abundance value in a given gel slice was then normalised to its maximum detection count, which means each protein has a relative abundance score that goes up to 1 along the gel. For visualisation, Excel conditional-formatting tool was used to change the colour of the cell according to the protein abundance, whereby a value 1 was displayed as red, 0.7 was displayed as yellow, 0 was displayed as black, and gradient for anything in-between.

2.10 Flavin characterisation assays

2.10.1 Preparation of FMN standards

FMN standards were prepared carefully using a metal weighing boat and volumetric flasks. 20 mg of FMN was weighed out and added to two separate 100 mL volumetric flask, and then the empty boat was weighed again, and the difference recorded. Thereafter, Milli-Q H_2O was added to 100 mL volume (with the bottom of the meniscus on the line) and the flask was shaken to dissolve the FMN. The two flasks were labelled A1 and B1 (~0.2 mg mL^{-1} of FMN). 2.5 g of solution from each of A1 and B1 were transferred into two 50 mL volumetric flasks (A10 and B10 respectively), and the flasks were filled with Milli-Q H_2O to make 10 µg mL^{-1} FMN standards. Standard B10 was exclusively used to verify the accuracy of the weighing of the FMN standards. A10 was further diluted (in size-exclusion buffer C without DDM) in 10 mL volumetric flasks to make 0.1, 0.2, 0.5, 1 and 2 µg mL^{-1} of FMN standards to be used in the assay. Standards were kept in the dark at 4 °C to prevent photo-degradation of FMN.

2.10.2 Quantification of FMN in complex I

10-16 µL of complex I (10 µg) and 10 µL of FMN standards (0.2 – 10 µg mL^{-1}) in size-exclusion buffer C without DDM were transferred into a 0.5 mL amber Eppendorf tubes. To each tube, 90 µL of 11% trichloroacetic acid (v/v) was added to extract FMN, and vortexed. After 15 min. of incubation at room temperature, the solution was centrifuged for 3 min. using a microfuge at maximum speed. 40 µL of each supernatant (containing FMN) was transferred into a microplate and then mixed with 200 µL potassium phosphate in the wells. The concentration FMN bound to complex I was estimated using the fluorescence intensity of known FMN standards (excitation at 450 nm and emission at 525 nm).

2.10.3 Thermo flavin assay

This assay was adapted from Forneris *et al.*, (Forneris et al., 2009). 20 µL of standards of 0.1 - 2 µg mL^{-1} of FMN (with size-exclusion buffer C without DDM as control) and complex I (0.3 µg mL^{-1}) were carefully added to an Applied Biosystems MicroAmp 96-Well reaction plate and sealed with AB MicroAmp clear adhesive film. The plate was inserted into an Applied Biosystems 7900HT PCR machine and calibrated with the FMN standards following the

manufacture's protocol for pure dyes. The machine was programmed to start at 20 °C for 2 min., and then increase the temperature by 1.5 °C every 30 sec. up to 99 °C. The fluorescence data was extracted in an Excel file, where the concentration of FMN bound to complex I was calculated from the known FMN standards, and the temperature in which 50% of FMN dissociates from complex I (T_{m50}) was calculated from a Boltzmann sigmoidal fit made in Prism.

2.11 EPR spectroscopy

John Wright from Queen Mary University of London carried out EPR spectroscopy measurements. Samples for electron paramagnetic resonance (EPR) were prepared under anaerobic conditions in a Braun UniLab-plus glovebox (O_2 < 0.5 ppm, N_2 atmosphere) at room temperature (22°C). 9 µL of complex I sample (10-20 mg mL^{-1}) was reduced with 2 mM NADH and rapidly transferred to a 1.6 O.D. suprasil quartz EPR tube (Wilmad, inner diameter 1.10 ± 0.05 mm). The sample was then flash-frozen by immersion in an ethanol cold-trap, which was cooled externally with a dry ice/acetone slurry. The samples were then transferred to liquid nitrogen and removed from the glovebox. EPR measurements were performed using an X/Q-band Bruker Elexsys E580 Spectrometer (Bruker BioSpin GmbH, Germany) equipped with a closed-cycle cryostat (Cryogenic Ltd, UK). All measurements were carried out at X-band using an ER 4118X-MS2 resonator (2 mm sample access). EPR measurement conditions were 3.2 mW power, 100 kHz modulation frequency and 0.7 mT modulation amplitude. Baseline measurements of the sample buffer were subtracted.

2.12 CryoEM grid preparation

Cryo-EM grids were prepared alongside Dr. James Blaza from the MRC-MBU, University of Cambridge. UltrAuFoil® gold grids (0.6/1, Quantifoil Micro Tools GmbH) were used to minimise electron beam induced motion during imaging (Russo and Passmore, 2014), and were prepared as described previously (Blaza et al., 2018). The grids (on a glass slide) were made

hydrophilic by glow discharging them at 20 mA for 60 sec. under vacuum with an EMITECH glow-discharger unit. The thiol-sensitive gold surface on the grids was then modified by incubation in 5 mM 11-mercaptoundecyl hexaethyleneglycol (SPT-0011P6, SensoPath Technologies) in ethanol for two days in an anaerobic glovebox (Meyerson et al., 2015). Following incubation, the grids were washed in ethanol and air-dried prior to use. Using an FEI Vitrobot IV and accompanying FEI equipment, 2.5 µL of complex I solution (3 – 5 mg mL^{-1}) were applied to each grid at 4 °C in 100% relative humidity and blotted for 8 – 16 sec. at a force setting of -10, before plunging the grids into liquid ethane.

2.13 Automated cryo-EM data acquisition

Automated data collections on grids for the active and de-active mouse compelx I datasets were carried out by Dr. James Blaza in Astbury Centre for Structural Molecular Biology, University of Leeds and by the staff at electron Bio-Imaging Centre (eBIC) - Diamond Light Source, Harwell Science and Innovation Campus in Oxfordshire, respectively. Images for the datasets were recorded with EPU software using an FEI 300 keV Titan Krios microscope. The datasets were collected in counting-mode, using a Gatan K2 detector and GIF Quantum energy filter in zero-energy-loss mode with a slit width of 20 eV, and a 100 µM (C2) objective aperture. Data for active complex I was imaged at 1.05 Å/pixel (nominally 1.07 Å/pixel, 47,600× actual magnification) and de-active complex I at 1.33 Å/pixel (nominally 1.30 Å/pixel, 37,600× actual magnification). The defocus ranges were -1.5 to -3.1 µm for the active mouse dataset, and -2.2 to -3.4 µm for de-active mouse complex I dataset, with the autofocus routine run every 10 µm. Data for the active mouse complex I were collected with a dose rate of 5 electrons $Å^{-2}$ $sec.^{-1}$, with a 10 sec. exposure captured in 25 frames. Data for de-active mouse complex I were collected with a dose rate of 3.25 electrons $Å^{-2}$ $sec.^{-1}$, with a 15 sec. exposure captured in 30 frames. The total dose was ~50 electrons $Å^{-2}$ for both datasets.

2.14 Image processing

Data for the de-active complex I was processed by Dr. James Blaza.

2.14.1 Hardware and software

Software requiring high computational demand was operated on a Linux (CentOS 6 operating system) High-Performance Computing (HPC) cluster comprised of several nodes with a dual-sledged, hyper-threaded (2x), 16-core central processing unit (CPU) and 256 GB of random access memory (RAM). For programs such as Gctf or MotionCor2, or other graphical processing unit (GPU)-accelerated applications (such as 3D classification), an 8-core hyper-threaded (2x) CPU and 256 GB of RAM nodes were used with graphical processing units (4 x GeForce GTX 1080 with 8 GB of RAM). All image processing steps were carried out in the RELION (version 2.1 patch b1) software suite (Fernandez-Leiro and Scheres, 2017; Kimanius et al., 2016).

2.14.2 Whole-frame motion correction and dose weighting

In-plane beam-induced horizontal movements of micrographs and radiation damage were modelled and corrected for in all of the datasets using MotionCor2 (version 1.0) (Zheng et al., 2017). Briefly, unaligned movie stacks in MRC format were processed using all available frames in 10 x 10 non-overlapping patches and 5 x 5 patches for de-active mouse complex I. A B-factor of 150 was applied to the micrographs. Dose-weighting was also performed at this stage with 1.6 electrons $Å^{-2}$ per frame on images from both datasets. No pre-exposure of radiation was accounted for. This produces motion-corrected, summed images suitable for further processing. Micrographs with excessive drift were discarded.

2.14.3 Estimating the contrast transfer function (CTF)

Defocus and astigmatism estimates were obtained for each micrograph using Gctf (version v1.06_sm_30_cu8.0_x86_64) (Zhang, 2016). The CTF of the non-dose weighted motion-corrected micrographs were estimated using the following constants: the spherical aberration of the microscope (2.7 mm), amplitude contrast of 0.08 and 0.1 for micrographs from the active mouse complex I and the de-active mouse complex I datasets, respectively. The resulting fitted

CTF models were compared with their experimental power spectra (intensity vs. spatial frequency). Any images displaying noticeable astigmatism (evidenced by elliptical power spectra) or crystalline ice contamination (an intense white ring in the power spectrum) were discarded.

2.14.4 Particle picking and extraction

Particle coordinates for de-active mouse complex I were selected manually in RELION from motion-corrected and CTF estimated micrographs. In contrast, particle coordinates for active mouse complex I were automatically selected in RELION, with 10 2D class averages of bovine complex I used as a reference (kindly provided by Dr. Hannah Bridges from the MRC-MBU, University of Cambridge). The references were low-passed filtered to 20 Å to avoid selection bias. In the case of the active mouse complex I dataset, the automatically selected coordinates were manually curated to discard non-protein (junk) particles.

Particles from selected coordinates were extracted in RELION with a box-size of 364 and 450 pixels for the de-active mouse complex I and the active mouse complex I datasets, respectively. The contrast of each particle was inverted and normalised.

2.14.5 2D Classification

To obtain 2D class averages free from non-complex I particles, extracted particles were subjected to reference-free 2D classification in RELION into 100 classes, employing a mask diameter of 340 Å. In some cases, such as during the processing of the active mouse complex I data, the 'ignore CTF's until first peak' function was utilised to more evenly distribute particles into all classes. Usually, several rounds of 2D classification were required to clean the datasets.

2.14.6 Movie refinement and particle polishing

Consensus EM maps were first generated using the '3D-auto refine' function in RELION, with the bovine cryo-EM map (EMD-2676) as a reference map. The bovine map was low-pass filtered to 60 Å to avoid model bias in the reconstruction. Following the generation of a map, movie refinement in RELION was carried out on a reduced number of frames for the active mouse complex I dataset in an effort to mitigate the effects of radiation damage in later frames

(Scheres, 2014), and preserve high-resolution structural information. Movie refinement was carried out on the active mouse complex I data with frames 1-12 out of 25 (total dose 24 electrons $Å^{-2}$) with a running average window of 3 frames. In contrast, all the frames (30 frames) from de-active mouse complex I data were used for movie refinement, with a 5 frame running average window. Following an initial movie refinement, the particles were subjected to 'Particle polishing" procedure in RELION to account for radiation damage and movement (from movie refinement), using default parameters for B-factor weighting (to improve particle single to noise ratio), and a soft-mask generated in the 'Mask create' function in RELION, initially from the bovine complex I model (5LC5).

2.14.7 3D classification and reconstruction of high-resolution cryo-EM maps

In later 3D classifications and high-resolution reconstructions (herein referred to as 3D refinement), a map generated from the dataset itself was used as a refinement reference in place of the initial bovine map (EMD-2676).

In the case of de-active mouse complex I, 3D classification was performed prior to particle polishing. Coarse-grained 3D classification with an angular sampling of 7.5° was used to discard particles that do not resemble intact complex I. Following particle polishing, data for de-active mouse complex I was subjected to a further round of 3D classification into 5 classes with an increasing angular sampling to 0.9°. The resolution of the reconstruction was limited to 8 Å in this case, to prevent overfitting. Homogenous de-active classes (maps) were selected based on how similar their cross-correlation coefficients (CCC) were to the known de-active bovine compex I structure (PDB accession code 5LDW) (Zhu et al., 2016a), by utilising the 'fitting' function in UCSF Chimera (Pettersen et al., 2004a). Data for active mouse complex I were subjected to coarse-grained 3D classification using an angular sampling to 3.7, into 5 classes. As with the de-active mouse complex I, classes were then selected based on their similarity to known complex I models using Chimera's 'fitting' function.

The final 3D-refinement for all complexes used solvent-flattened Fourier shell correlation (FSC) curves, using masks to cut the volume of solvent considered and increase the orientation accuracy. In order to reduce the interference of the disordered detergent belt on resolution estimation, post-processing was performed in RELION, with default parameters using masks created from fitted coordinate files (using the Molmap command in Chimera and mask creation

tools in RELION). The resolution estimates were based on the gold standard FSC = 0.143 criterion (Scheres and Chen, 2012).

2.15 Model building

Most of the model building for the mouse complex I structures was carried out by Dr. James Blaza. The model for active mouse complex I was built in COOT (Emsley et al., 2010), using the published model of active bovine complex I (PDB accession code 5LC5) as a template, with the sequences converted to the mouse sequences. The atomic co-ordinates were initially adjusted into the electron density using local real-space refinement in COOT. Lipids were built and fitted into unoccupied density. The nucleotide bound to NDUFA10 was tentatively built as ADP because density indicates the presence of a purine ring and two phosphates and ADP is highly abundant in the mitochondrial matrix. Post-translational modifications were built using data from molecular mass measurements of mouse complex I subunits (Chapter 3.10), such as acetylation and hydroxylation of Arg-85. The entire model was then refined against the map by real-space refinement in PHENIX (Afonine et al., 2013). The model was then locally refined against the map in COOT. Multiple cycles of global refinement, followed by local refinement were assessed by MolProbity (Williams et al., 2017). The de-active mouse complex I, was built using the final active mouse complex I structure as a template.

2.16 Bioinformatic analyses

Global structure resolutions were estimated from FSC curves in RELION (Scheres and Chen, 2012), whilst local resolutions were estimated using the Local Resolution function in RELION with B-factor setting the same as that obtained in the post-processing step, and with default parameters used (sphere of radius 12.5 Å with a soft edge of 25 Å). For individual subunits, Chimera was used to extract the resolution at each atomic coordinate and the values were averaged. Values calculated by masking individual subunit volumes in the local resolution map gave equivalent results.

Secondary structures were analysed using DSSP (http://swift.cmbi.ru.nl/gv/dssp/) (Touw et al., 2015). Surface areas of contacts between subunits were calculated using the PDBePISA server (http://www.ebi.ac.uk/pdbe/pisa/) (Krissinel and Henrick, 2007). The ubiquinone-binding channel was predicted using CAVER 3.0 (Petrek et al., 2006), with a 1.4 Å probe radius, starting from the sidechain oxygen atom of NDUFS2-Tyr108. Protein structure and density figures were created using Pymol or Chimera.

Chapter 3: Purification and characterisation of complex I from *Mus musculus*

3.1 Introduction

3.1.1 Mouse as a model organism for complex I studies

In order to carry out effective structural and functional studies on complex I, it is imperative to first pick a suitable model organism. Currently, the model organisms used for complex I studies range from bacteria such as *E. coli* to mammals such as *B. taurus* (bovine). Bacterial models such as *E. coli* and *T. thermophilus*, in which the early structural work on complex I was carried out (Baradaran et al., 2013; Efremov and Sazanov, 2011), have the advantage that genetic manipulation can be carried out with relative ease allowing for mutation(s) in the enzyme to probe critical residues involved in its function. Although genetically amenable, and with yields sufficient for mechanistic studies, the bacterial enzyme is a distant relative to its mammalian counterpart. The mammalian enzyme has an additional 31 supernumerary subunits on top of the 14 shared core subunits: therefore, studies on assembly and dysfunction by way of pathogenic mutations in the supernumerary subunits are precluded by their absence in bacteria.

The yeast *Y. lipolytica* serves as a more relevant model because its enzyme is more closely related to the human enzyme than the bacterial enzyme (Abdrakhmanova et al., 2004). Additionally, the nuclear genome can be genetically manipulated, and yields sufficient complex I to allow for structural and functional studies (Abdrakhmanova et al., 2006; Zickermann et al., 2015), and although the yeast enzyme is a good model to study complex I function and assembly, its use in a mitochondrial disease context is limited because induced mutations in yeast based on known pathogenic mutations in humans cause phenotypes that fail to fully recapitulate the effect seen in the human enzyme (Varghese et al., 2015).

The bovine enzyme, due to its high sequence homology to the human enzyme, has provided a plethora of information on function, structure and subunit composition that are physiologically and pathophysiologically relevant to the human enzyme (Hirst et al., 2003; Kussmaul and Hirst, 2006; Zhu et al., 2016a). Moreover, greater yields of complex I are easily attainable using bovine heart tissue over bacteria and yeast. However, unlike in the bacteria and yeast models,

it is not feasible to genetically manipulate the bovine enzyme, therefore preventing studies on clinically or catalytically-relevant mutations.

With the use of immortalised human cell lines, research on complex I assembly, amongst other things, has been performed on cancer cells (Stroud et al., 2016). A limitation of working with human cells is that they have abnormal metabolism and therefore do not reflect normal physiology or respiration. Secondly, due to the limitation of material, structural and functional work with cell line complex I will prove very difficult.

A good model that fulfils most of the positive criteria discussed here is mouse. First, mouse heart is densely packed with mitochondria and can be prepared in sufficient yields and high activity to allow for comprehensive studies (Fernández-Vizarra et al., 2010; Fernández-Vizarra et al., 2002). Secondly, unlike bovine and other related large mammalian species, mouse is more genetically amenable, with available complex I deficiency models, which recapitulate some of the clinical phenotypes seen in patients (Kruse et al., 2008). A list of some of the complex I deficiency models in mouse are shown below (Table 3.1). With this in mind, mouse was explored as a suitable model for complex I studies.

Gene	Genetic manipulation	Phenotype	Reference
Ndufs4	Whole body knockout	Leigh-like syndrome	(Kruse et al., 2008)
Ndufs6	Heart-specific knockdown	Cardiomyopathy	(Ke et al., 2012)
Ndufa5	Neuron-specific conditional knockout	Mild chronic encephalopathy	(Peralta et al., 2014)
Ndufa13	Heart specific conditional knockout	Cardiomyopathy	(Hu et al., 2017)
Nd6	Homoplasmic G13997A mutation: P25L	Leber hereditary optic neuropathy	(Lin et al., 2012)

Table 3.1: Available complex I disease models with the corresponding genetic manipulation and observed phenotype. The table was adapted from the review of Diaz and co-workers (Torraco et al., 2015).

3.1.2 A brief historical perspective on purification of mammalian complex I

Preparation of bovine complex I was first achieved in 1962 using an ammonium sulphate precipitation method (Hatefi et al., 1962). The preparation displayed modest inhibitor-sensitive activity and retained high amounts of phospholipids, despite the harsh treatment of the enzyme.

Moreover, further characterisation showed 1 FMN and 26 non-haem irons per enzyme; cofactors required for NADH:UQ_{10} oxidoreduction. However, using spectroscopic techniques on the preparation identified cytochromes as a significant component in the preparation. In addition, the purification method resulted in complex I being polydisperse and therefore not suitable for structural studies. More than three decades later, the first chromatographic method for purifying complex I was developed by Walker and co-workers (Walker et al., 1995). In brief, membranes were solubilised with 1% DDM. Then, complex I was enriched by ammonium sulphate precipitation. Finally, complex I was further purified using an anion-exchange column. Although the enzyme was pure, it displayed low NADH:Q_1 activity that was rotenone insensitive. Further analysis of complex I by SDS-PAGE showed that NDUFA10 was substoichiometric compared to the Hatefi preparation (Walker et al., 1995). Complex I prepared by this method was used to identify its subunit composition by mass spectrometry. Subsequently, Walker and co-workers attempted to modify the protocol, by co-purifying complex I and V (Buchanan and Walker, 1996). For the first time, a size-exclusion column was used, initially to strip phospholipids off the enzyme (for crystallisation trials) and desalt the preparation following ammonium sulphate treatment. Then, two anion-exchange columns and a final size-exclusion column further resolved the preparation. Although complex I was pure and monodisperse, it again displayed low activity (but rotenone sensitive). As with previous attempts to purify complex I, NDUFA10 was still being lost during purification. Later attempts to purify complex I by Walker and co-workers (Sazanov et al., 2000), by maintaining the ammonium sulphate precipitation step, but using two successive anion-exchange column and a size-exchange column resulted in a pure complex I, however, it was devoid of lipids and showed low catalytic activity.

As phospholipids were shown to be crucial for complex I catalytic activity (Buchanan and Walker, 1996; Sazanov and Walker, 2000; Walker et al., 1995), Hirst and co-workers prepared the enzyme with various phospholipid mixtures (Sharpley et al., 2006). They showed that by supplementing soybean asolectin (phospholipid mixture) in the anion-exchange buffer, a higher activity was recorded than the preparation without asolectin added. Complex I purified using this method was pure, monodisperse and was highly active. Furthermore, lipid analyses showed that the preparation contained appreciable amounts phosphatidylcholine, phosphatidylethanolamine and cardiolipin, reflective of the relative phospholipid composition of the inner mitochondrial membrane. Moreover, the length of the protocol was reduced significantly from previous purifications, by utilising a single anion-exchange column (instead

of two used in the Walker and co-workers preparations), followed by a size-exclusion column. Compared to previous bovine complex I preparations, NDUFA10 stoichiometry was increased in this purification. This purification methodology served as the foundation for future complex I preparation that was essential for structural and mechanistic studies (Jones et al., 2016; Vinothkumar et al., 2014; Zhu et al., 2016a).

Recently, complex I was purified from ovine as an alternative model to bovine for structural studies (Letts et al., 2016b). In essence, the purification followed that of the bovine enzyme (Sharpley et al., 2006), with a single anion-exchange step, followed by size exclusion chromatography. The enzyme preparation differed in that LMNG was used to extract the complexes from membranes as opposed traditional DDM used in bovine. The authors determined the suitability of using LMNG over other detergents by monitoring NADH:DQ (a quinone analogue) activities by incubating membranes in various detergents. In the ovine complex I purification, lipids were supplemented in the size-exclusion buffer rather than the anion-exchange buffer (as done previously (Sharpley et al., 2006), where it is known that lipids are lost in the anion-exchange step). The resulting complex I preparation although pure, displayed low catalytic activity in comparison to the recent bovine enzyme preparation (Jones et al., 2016). Although the detergent Brij-35 was shown to substantially decrease NADH:DQ activity in membranes, it was used for the structural studies of ovine complex I (Fiedorczuk et al., 2016).

By using previous bovine work as the bases for purification, a scaled-down protocol was developed to purify complex I from five different rat tissues (Mohammed PhD thesis, 2015). Due to issues of protein stability with DDM, LMNG was explored as an alternative detergent for extracting the complexes from membranes. The chromatographic preparation following solubilisation was the same as bovine, but with an additional affinity step, utilising ATPase inhibitor (IF_1) in the column matrix to resolve ATPase from complex I. The preparation was relatively pure but had low complex I activity making it impractical for either structural or mechanistic studies (Mohammed PhD thesis, 2015).

3.2 Aims of this chapter

The work in this chapter aims to:

- Develop a protocol to prepare mouse complex I in a pure, stable and active enzyme, which can serve as a basis for structural, functional and clinically relevant studies.
- Comprehensively characterise the enzyme in terms of kinetics, protein content and spectroscopic properties.

3.3 Catalysis by mouse mitochondrial membranes

Prior to complex I purification, coupled mitochondria were isolated from mouse hearts using an adapted differential centrifugation method (Fernández-Vizarra et al., 2010; Fernández-Vizarra et al., 2002). Briefly, Mitochondria were burst open using high-energy sonication, followed by ultra-centrifugation to obtain membranes for further experiments, as described in Chapter 2.4 To assess the suitability of mouse mitochondrial membranes for studies on oxidative phosphorylation enzymes, kinetic assays were carried out (see Chapter 2.6) and compared to the current and well-characterised bovine system (Table 3.2). The NADH:O_2 rate, which is a measure of catalysis by complexes I to IV, revealed a higher activity for mouse membranes than bovine membranes. Plausibly, the higher activity may be attributed to the shorter preparation time, leading to better preservation of the enzymes (Fernández-Vizarra et al., 2010).

Assay	Mouse (μmol min^{-1} mg^{-1})	Bovine (μmol min^{-1} mg^{-1})
NADH:O_2	1.47 ± 0.08	0.80 ± 0.02
Succinate:O_2	1.22 ± 0.02	0.86 ± 0.07*
ATP hydrolysis	3.82 ± 0.07	2.28 ± 0.02*
H_2O_2	$3.53\times10^{-3} \pm 0.12\times10^{-3}$	$5.90\times10^{-3} \pm 0.20\times10^{-3}$†
FeCN	5.13 ± 0.07	n.d.
HAR	5.71 ± 0.06	4.63 ± 0.20
APAD	0.75 ± 0.02	n.d.

Table 3.2: Enzymatic activity measurements on mitochondrial membranes. Enzymatic activities were measured according to the protocols described in Chapter 2.6 unless stated otherwise. *Succinate:O_2 and ATP hydrolysis by bovine samples were measured in submitochondrial particles (SMPs) rather than membranes; 2.5 µg mL^{-1} gramicidin was added to collapse the proton motive force. † Rate of H_2O_2 production by bovine SMPs was obtained from Hirst and co-workers (Pryde and Hirst, 2011). n.d. (i.e. not determined). The data are presented as mean values with error bars for the standard error of the mean (n ≥ 3).

Likewise, the stoichiometric and spectrophotometrically sensitive succinate:O_2 assay (complexes II to IV) displayed higher activity for mouse membranes than bovine membranes, albeit it has been previously reported that bovine samples could reach similar rates (Jones and Hirst, 2013). ATPase (complex V), a canonical enzyme was tested by ATP hydrolysis. Although the activity was higher for mouse membranes than bovine membranes, it is cautioned that SMPs and membranes may differ in the relative amount of the proteins present being measured (Pryde and Hirst, 2011). To establish NADH oxidation activity by complex I in membranes, the flavin-site was probed by artificial oxidants with differing modes of action (Birrell et al., 2011). Although the rates of FeCN and APAD reduction (by ping-pong mechanisms) were measured in mouse, the equivalent bovine data is not available. Nevertheless, the rate of reduction of HAR (by a ternary mechanism) is known for both species and is broadly consistent. Lastly, the rate of reactive oxygen species formation was measured by monitoring the oxidation of Amplex Red to resorufin, in the presence of NADH. H_2O_2 production rates by mouse membranes were almost half those of bovine membranes. As the membranes of both species contain a complicated mixture of proteins, and possible extra-mitochondrial protein contamination, not all of the reactive oxygen species can be confidently attributed to complex I. The higher activities (oxidative phosphorylation enzymes) exhibited by mouse membranes compared to the already highly active bovine membranes, suggests that mouse membranes are suitable for complex I preparation.

3.4 Establishing optimal solubilisation conditions for mouse complex I

It was critical to establish a solubilisation condition to maximise the amount of complex I yielded by the membranes without compromising stability or causing protein aggregation. BN-PAGE, SDS-PAGE Western-blot and complex I-specific activity measurements were used to assess different detergents and their concentrations. Based on previous work with complex I from rat, for which detergent-based issues lead to protein instability and aggregation (Mohammed PhD thesis, 2015), mouse membranes were incubated with varying concentrations of a mild detergent; LMNG. Following the incubation, stability, yield and activity of complex I were assessed (Figure 3.1). BN-PAGE analysis revealed that, although complex I seemed to be stable in LMNG, even at higher concentrations, there was also a band with a higher apparent molecular weight than monomeric complex I, with equal staining intensity with the redox-

sensitive NBT dye (Figure 3.1a). This may indicate a higher-order molecular organisation that persists under the mild detergent (i.e. a supercomplex) or possible dimerisation of the enzyme. To confirm this higher band observed on BN-PAGE is complex I, and not an alternative flavin-bound protein that is being solubilised using LMNG, soluble and pellet fractions resulting from solubilisation (in various detergent concentration) and centrifugation of membranes, were separated on SDS-PAGE and analysed by Western blot. Two subunits, present at opposite ends of the complex, were investigated (NDUFAB9 and NDUFS1) (Figure 3.1b). As illustrated by SDS-PAGE Western blot analyses, an increase in the LMNG concentration is followed by an increase in the presence of complex I subunits detected in the soluble fractions and an accompanying decrease of those very same subunits in the pellet fractions. Confirmatory to this, the complex I-specific activity in the solubilised fractions increased with increasing detergent concentration. Therefore, the majority of complex I was not solubilised in its monomeric form, making LMNG a non-ideal choice of detergent for purifying monomeric complex I. DDM, a mild detergent often used in solubilisation of membranes, which is reported to be apparently less stabilising than LMNG, but that can break inter-complex interactions, was thus explored (Figure 3.1) (Chae et al., 2010). BN-PAGE analyses show there are no visible breakdown products resulting from the solubilisation, and that with more than 1% DDM concentration, complex I exists mostly as a monomer (Figure 3.1a). As with LMNG, SDS-PAGE Western blot shows an increase in detergent concentration led to an increase in the complex I subunits detected in the solubilised fractions (Figure 3.1b). In addition, beyond 0.5% detergent concentration, total complex I activity in the solubilised fractions were greater for DDM than LMNG, correlating to a greater yield for DDM. Careful consideration of both detergents suggests that DDM is superior to LMNG by both increasing the yield of complex I obtained from membranes and solubilising complex I as a monomer without having an adverse effect on the stability of the enzyme. After careful consideration, DDM was deemed a suitable choice of detergent for purification and was used at 1% concentration to solubilise mouse membranes.

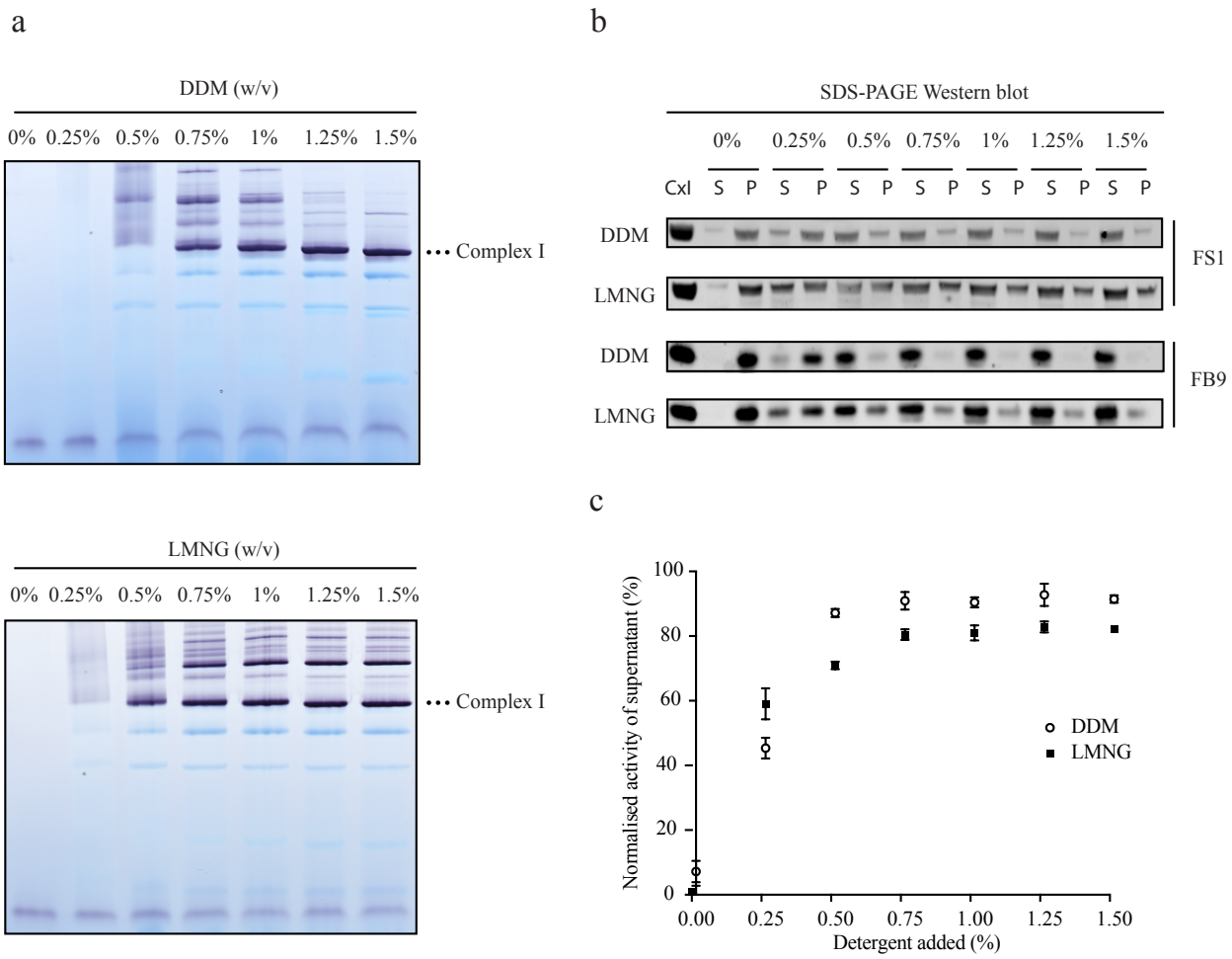


Figure 3.1: Evaluating solubilisation conditions on mitochondrial membranes. Visualisation and assays were carried out with an equal volume of material (i.e. the pellet was re-suspended to the same volume as the supernatant). (a) Solubilised membranes were analysed by BN-PAGE and monomeric complex I (indicated as complex I) stained with 0.5 mg mL^{-1} NBT and 100 µM NADH. (b) Solubilised and pellet fractions from the solubilisation process, were subjected to SDS-PAGE and Western blotting by incubation with anti-NDUFS1 (FS1) and anti-NDUFB9 (FB9) antibodies. Stained proteins were imaged using a LI-COR imaging system. (c) The NADH:HAR activity measurements using 100 µM NADH and 3.5 mM HAR were carried out on supernatant and pellet fractions, with rates of supernatants (%) normalised to the sum of both rates (i.e. pellet and supernatant). The data are presented as mean values with error bars for the standard error of the mean (n ≥ 3).

3.5 Elution of complex I using an anion-exchange column

Anion-exchange purification relies on the negatively charged surface of a protein binding to the positively charged groups present on the column matrix. The change in the surface charge of a protein is in turn modulated by the pH of the buffer, through the protonation and de-protonation of solvent-exposed amino acid side chains. Despite sharing high sequence homology with its bovine counterpart, slight changes in the sequences of the solvent-exposed amino acids in mouse complex I may alter the surface charge properties of the protein, thus affecting its binding property and ultimately, its elution from the column. Initially, solubilised membrane proteins were loaded onto a HiTrap^{TM} Q HP column, pre-equilibrated with buffer containing a lower concentration of DDM (0.1%) than previous bovine preparations (0.2%) to reduce complex I delipidation, and supplemented with an asolectin (phospholipid mixture) from soybean to further minimise delipidation of the enzyme (Sharpley et al., 2006). To establish the elution profile of the enzyme, the salt concentration (NaCl) was adjusted in a stepwise manner and protein elution monitored using an Äkta micro system (Figure 3.2a). Five protein-containing peaks eluted from the column in various salt concentrations, with the highest peak eluting at ~320 mM NaCl. The first peak (~ 1 mL) was not considered as it lacked colouration and was therefore unlikely to contain complex I. NADH:HAR measurements on the remaining eluted peaks showed that complex I mainly eluted at ~320 mM NaCl, with residual amounts of the protein found elsewhere (Figure 3.2b). Complex I from mouse thus elutes from the column in an equivalent concentration of salt to the bovine enzyme (Sharpley et al., 2006). In light of the obtained elution profile (Figure 3.2), in later preparations, complex I-enriched fractions were eluted in 350 mM NaCl.

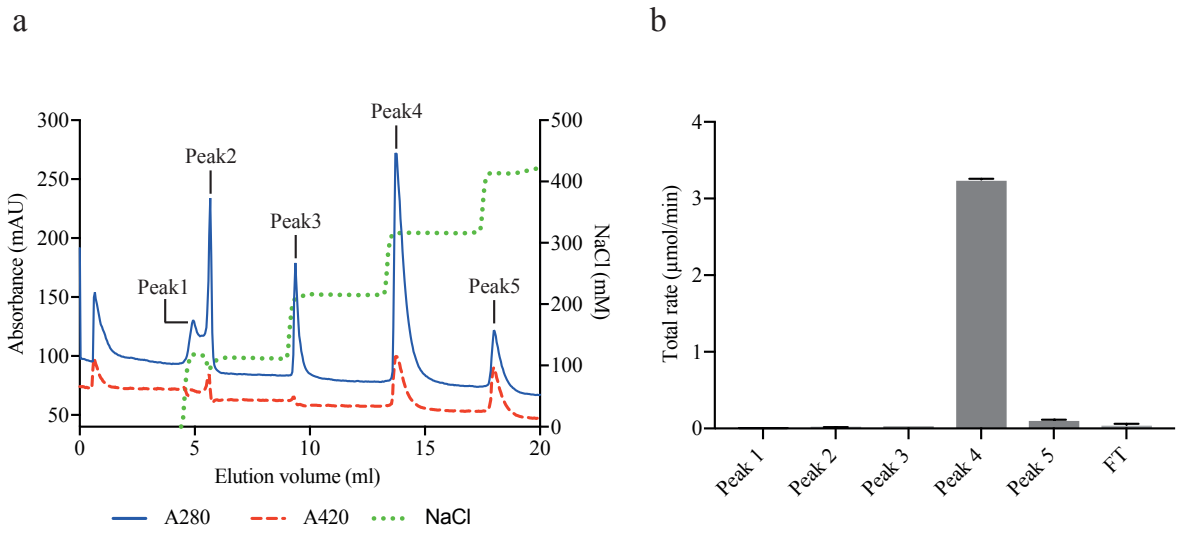


Figure 3.2: Complex I elution profile from an anion-exchange column. The Äkta micro system and the attached 1 mL HiTrap™ Q HP column were washed and equilibrated at a flow rate of 0.3 mL min^{-1} with buffer A (20 mM Tris-HCl pH7.14 at 23 °C, 1 mM EDTA, 0.1% DDM, 10% ethylene glycol, 0.005% asolectin and 0.005% CHAPS), and proteins were eluted with buffer B (buffer A + 1 M NaCl). (a) Solubilised membranes (1 mg) were separated with an increasing concentration of NaCl (green broken line), and the eluents absorbance measured at 280 nm (blue solid line) and 420 nm (red broken line). (b) NADH:HAR activity measurements were carried out using 5 µl of each peak fraction with 100 µM NADH and 3.5 mM HAR. The total activity was calculated by taking into account the total volume in each fraction. The data are presented as mean values of technical replicates with error bars for the standard error of the mean ($n \geq 3$).

3.6 Isolation of pure and highly active complex I

Although anion-exchange purification yields a complex I-enriched fraction, it does not guarantee that the eluent will be pure, because proteins of different masses but with similar charges will elute together. Thus a size-exclusion column was utilised to further purify complex I in a size-dependent manner (Figure 3.3). Moreover, the size-exclusion column also behaves like a buffer exchange system, ideal for decreasing both salt and detergent concentration for future structural work. Complex I elutes from the size–exclusion column as a monodisperse peak at ~1.36 mL, followed by complex IV at ~1.65 mL (Figure 3.3a), with minimal overlap

between the two peaks. In addition, proteins at the intersecting regions were excluded when the complex I-containing fractions were pooled and concentrated, avoiding possible contamination. Following purification, complex I-containing fractions from each step were further inspected by BN-PAGE to visually assess their purity (Figure 3.3b). The complex I sensitive dye shows that during purification, the purity of the sample increases with each step, accompanied by a noticeable decrease in the level of contaminating proteins, resulting in a visibly pure enzyme that can be further characterised. To accurately quantify the recovery of complex I during the purification procedure, the NADH:DQ rates of each purification step fraction (containing complex I) was measured (Table 3.3). In agreement with the BN-PAGE analyses, with purification, there is an enrichment of complex I, as illustrated by an increase in specific activity. Although the preparation seems pure and highly active, in comparison to previous mammalian preparations (Letts et al., 2016b), the final complex I yield recorded was ~16%, by normalising the total activity of the purified sample to the total activity measured in membranes. The low yield could be explained by considering that during purification, a significant amount of each fraction was removed to ensure accurate estimation of protein content and to measure NADH:DQ rates. Taking into account that the starting material was only 2.5 mg of protein, removing fractions for analyses affects the yield of complex I substantially. In later preparations that have not been so heavily analysed, the complex I recovered was ~30 - 40%, based on total complex I content in the membranes used for purification. Bearing in mind that the starting material for bovine complex I preparations (~2 kg per heart) is ×10,000 greater than for mouse (~0.2 g per heart), this purification has facilitated the efficient recovery of pure and highly active mouse complex I, despite the relatively limited amount of starting material.

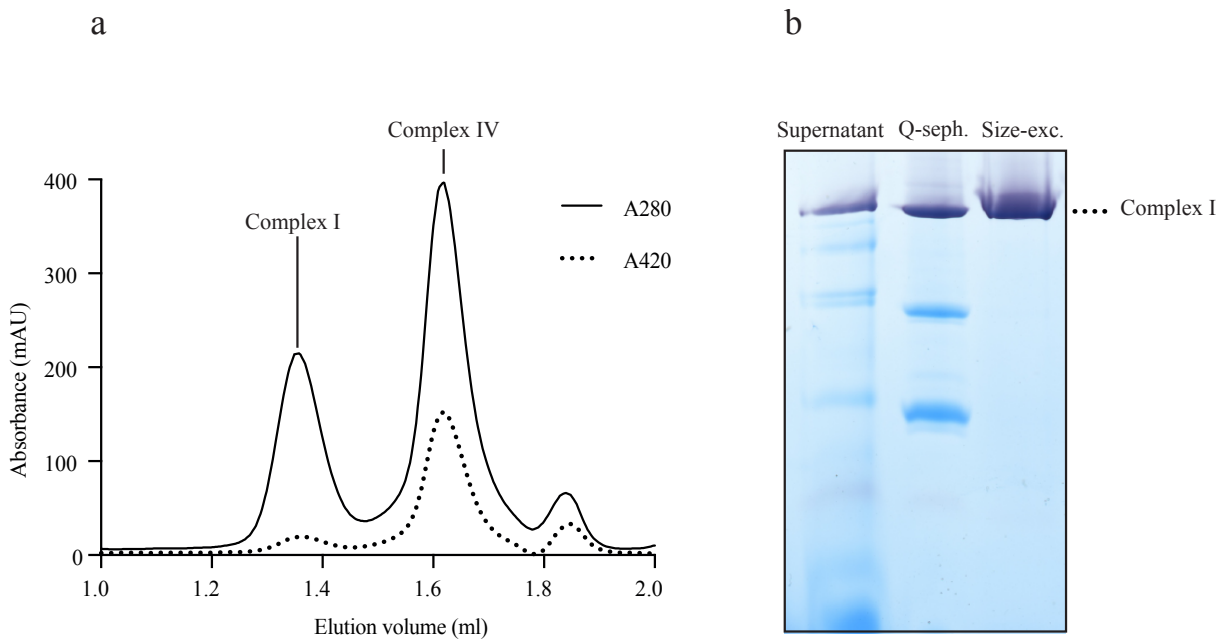


Figure 3.3: Chromatographic purification and visualisation of complex I. The Äkta micro system was operated at a flow rate of 0.03 mL min^{-1} with buffer C (20 mM Tris-HCl pH 7.14 at 23 °C, 10% ethylene glycol, 150 mM NaCl and 0.02% DDM), using a Superose 6^{TM} Increase 3.2/300 column (a) Pooled and concentrated Q-sepharose fractions were loaded and proteins were separated by size-exclusion chromatography, resulting in a size-based separation of complex IV and I. (b) Collected fractions: size-exc. (5 µg), Q-seph. (10 µg) and supernatant (10 µg) were visualised on BN-PAGE by staining with 0.5 mg mL NBT and 100 µM NADH.

Fraction	Total activity (µmol min^{-1})	Total protein (µg)	Specific activity (µmol min^{-1} mg^{-1})	Activity recovered (%)
Membrane	2.96 ± 0.03	2500	1.18 ± 0.01	100
Supernatant	2.06 ± 0.06	2070	1.00 ± 0.03	70
Q-sepharose	0.64 ± 0.02	234	2.75 ± 0.09	22
Size-exclusion	0.47 ± 0.01	46	10.38 ± 0.25	16

Table 3.3: NADH:DQ activity measurements on the fractions from a mouse complex I preparation. Rates for membranes (3 µg mL^{-1}), solubilised supernatant (3 µg mL^{-1}), Q-sepharose fraction (3 µg mL^{-1}) and complex I (0.5 µg mL^{-1}) were measured with 100 µM DQ in 0.075% (w/v) asolectin, 0.075% (w/v) CHAPS and 100 µM NADH, in 20 mM Tris-HCl pH 7.55 at 32 °C. The data are presented as mean values with error bars for the standard error of the mean (n ≥ 3).

3.7 Activity measurements on purified mouse complex I

To validate mouse complex I as a suitable model for future work, it was necessary to kinetically characterise and compare the enzyme to the well-established bovine model. The rates of reactive oxygen species production, flavin-site reactions and reduction of a quinone analogue (DQ) were investigated (Table 3.4). The catalytic rate of NADH:DQ oxidoreduction is twice as high in bovine as in the mouse. In later preparations of mouse complex I, a superior size-exclusion was used (Superose 6^{TM} Increase 5/150 column) to reduce the time during purification, by using a faster flow rate. NADH:DQ activity measurements from this size-exclusion column reached ~15 μmol min^{-1} mg^{-1} (Yin, personal communication). Nevertheless, the mouse enzyme exhibits a far higher activity than the current ovine complex I preparation (Letts et al., 2016b), which may be attributed to the supplementation of lipids early on in the preparation (Sharpley et al., 2006). Looking at the H_2O_2 production between the two species there is an appreciable difference in the level of reactive oxygen species produced, with complex I from mouse having a lower activity than its bovine counterpart. The reason for mouse displaying a low rate of reactive oxygen species was not explored, due to the limitations from the scarcity of the material. For example, it would be necessary to vary the total nucleotide concentration and adjust the NADH/NAD$^+$ ratio, to better characterise the reaction by altering the rates observed (Kussmaul and Hirst, 2006), and therefore these experiments would require large amounts of enzyme. Interestingly, the difference in the rates observed here, under similar conditions, may underlie a subtle physiological or pathophysiological difference in superoxide production by the two species. Finally, the flavin-site activities, with artificial flavin oxidising substrates that display both ping-pong (FeCN) and ternary (HAR) mechanisms were compared (Birrell et al., 2009). Both the HAR and FeCN rate was significantly greater for bovine than mouse. This is strange, considering that NDUFV1 (i.e. site of flavin reaction) is highly conserved (96% sequence identity) between the two species. Although this difference in flavin-site activities is not yet clearly understood, the mouse enzyme displays high rate of physiologically-relevant catalysis (NADH:DQ), making it suitable model for further characterisations.

Assay	Mouse ($\mu mol\ min^{-1}\ mg^{-1}$)	Bovine ($\mu mol\ min^{-1}\ mg^{-1}$)
NADH:DQ	11.99 ± 0.11	~24*
H_2O_2	15.30 $\times 10^{-3}$ ± 0.19	21.10 $\times 10^{-3}$ ± 2.90†
FeCN	32.87 ± 1.74	~72‡
HAR	47.38 ± 1.49	~90‡
APAD	4.46 ± 0.09	-

Table 3.4: Kinetic characterisation and comparison of purified complexes I. Rates for the purified enzyme were measured according to the protocols in chapter 2.6 * NADH:DQ activity was measured on purified bovine complex I (0.5 $\mu g\ mL^{-1}$) using 200 µM NADH and 200 µM DQ, instead of the 100 µM NADH and 100 µM DQ used for purified mouse complex I (Blaza et al., 2018). † H_2O_2 production was measured on purified complex I with 30 µM NADH, 2 units mL^{-1} HRP and 10 µM Amplex Red (Kussmaul and Hirst, 2006). ‡ Rate measurements were carried on purified complex I with 100 µM NADH, 1 mM FeCN and 3.5 mM HAR (Birrell et al., 2009). The data are presented as mean values of technical replicates with the error represented as the standard error of the mean ($n \geq 3$).

3.8 Estimating the flavin content of purified mouse complex I

The flavin-site reactions described in (Table 3.4) represent the initial step of the proton-coupled NADH:UQ_{10} oxidoreduction reaction and require a tightly bound FMN to function properly. As mentioned previously, there were significant differences in the flavin-site activities between mouse and bovine complex I (Table 3.4). To investigate whether these changes were due to complex I losing flavin during purification, the FMN content of purified mouse complex I was measured by reference to a set of prepared FMN standards (Figure 3.4). The flavin bound to complex I was found to be essentially stoichiometric relative to the total enzyme content (Figure 3.4a), therefore the low flavin-site activities were not due to loss of flavin during purification. In addition, in the same assay, the release of flavin bound to complex I upon increasing temperature was monitored (Figure 3.4b). The assay works by measuring the emission of photo-excited flavin. Under native conditions, the surrounding protein absorbs the energy from the excited flavin, quenching the measured signal. Under denaturing conditions

(i.e. increased temperature), the flavin is released, and once solvent exposed, the emission is unimpeded and is reflected by an increase in fluorescence intensity. Identifying the point at which 50% of the FMN is detectable (Tm_{50}) is important in understanding the tightness of the association between complex I and FMN, and allows for interspecies comparison, and comparisons between wild type and disease-related models. Across a wide selection of organisms, complex I Tm_{50} values for FMN range from 42.7 °C in *Pichia pastoris* to 94.7 °C in *T. thermophilus* to 60.1 °C in bovine (Bridges and Hirst, unpublished). The bovine enzyme binds significantly more tightly to FMN than its mouse counterpart (Tm_{50}=51.23) (Figure 3.4). It is currently unclear what the relationship is between the binding affinity of FMN to complex I and its function.

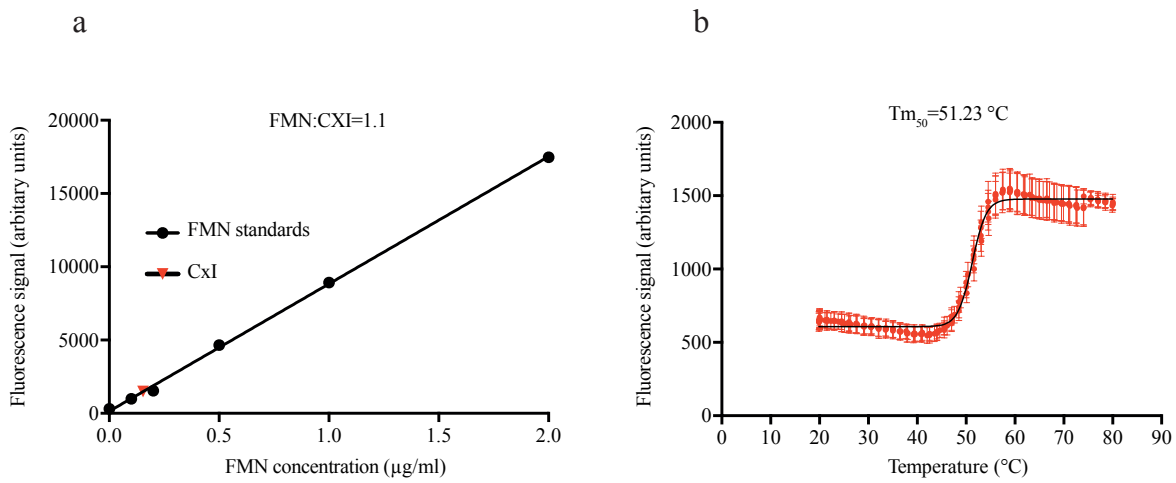
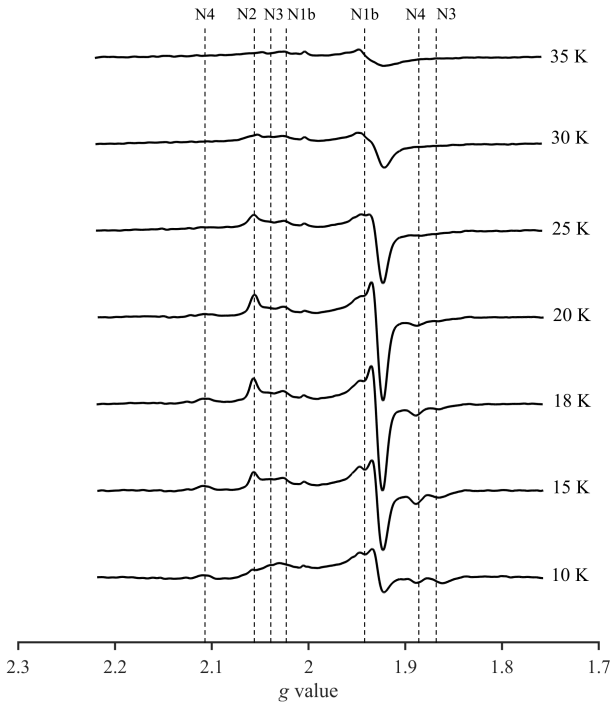


Figure 3.4: Flavin content and binding properties of mouse complex I. Details of the methods employed are described in Chapter 2.10 (a) Flavin quantification: The concentration of FMN bound (closed red triangle) to purified complex I (~6 µg) was estimated by reference to a set of known standards (closed circle) ranging from 0.1 – 2.0 µg mL^{-1}. The number of FMN per complex I (FMN:CXI) was calculated and displayed (n=3). (b) Flavin binding: purified complex I was exposed to increasing temperature from 20 – 80 °C, and the fluorescence signal of FMN was measured. Tm_{50} is the temperature at which 50% FMN is detectable measured using the fitted model (black sigmoidal curve).

3.9 EPR measurements on the Iron-sulphur clusters in mouse complex I

A crucial feature in the function of complex I is the ability to transfer electrons from NADH to ubiquinone, across the peripheral arm. To facilitate efficient electron transfer a chain of iron-sulphur clusters are embedded in the protein, within favourable electron transfer distance of one another (Page et al., 2003). In order to observe and characterise the iron-sulphur clusters, continuous wave X-band EPR spectroscopy was employed. In essence, EPR looks at how unpaired electrons in co-factors or radical amino acid side chains behave when a changing external magnetic field and a fixed microwave frequency is applied. When the magnetic field-induced energy gap between the two spin states resonates with the microwave frequency, the unpaired electron jumps to a higher energy state and subsequently relaxes, resulting in microwave absorbance. To measure EPR spectra for the iron-sulphur clusters, complex I was reduced with NADH (oxidised 2Fe and 4Fe clusters are diamagnetic and therefore are invisible in EPR spectroscopy), and immediately frozen. EPR signals were readily assigned to known clusters based on previous work on the bovine enzyme, and their positions are indicated in the structure (Figure 3.5a, b) (Wright et al., 2016). As with the bovine enzyme, the signals observed for the iron-sulphur clusters are temperature dependent; for example, N2 cluster has a maximum signal detected between 18 - 20 K, whilst N1b, N4 and N3 have a maximum signal between 10 – 15 K. Present in the bovine enzyme, but absent from the mouse enzyme, is the signal from cluster N5, which is only observable at 4 K because it is fast relaxing, (Reda et al., 2008a). Also, as with the bovine enzyme, not all of the iron-sulphur clusters present in mouse complex I were detected (4 out of 8). Using Mössbauer spectroscopy, Hirst and co-workers showed that the EPR-silent clusters remain oxidised when the enzyme is reduced by NADH (Bridges et al., 2012). Although the chain of clusters is partially oxidised, the first cluster to accept electrons from reduced FMN (N3) and the terminal cluster (N2) involved in quinone reduction, both show up clearly in the spectra suggesting the chain of iron-sulphur clusters is fully functional. EPR data shows a close match between the mouse enzyme and the highly characterised bovine enzyme (Wright et al., 2016).

a



b

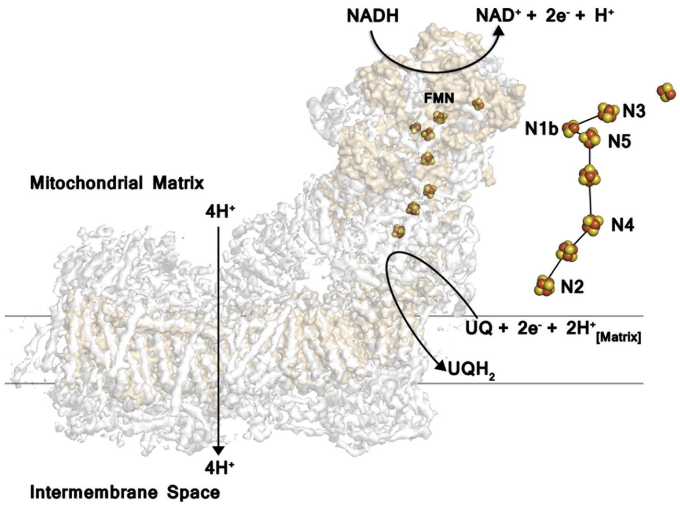


Figure 3.5: X-band continuous wave EPR spectra of mouse complex I. (a) Detailed conditions of the EPR measurements are given in Chapter 2.11. 9 µl of complex I (180 µg) were reduced by 2 mM NADH and the first derivative of the absorbance of the reduced iron-sulphur clusters (labelled) were detected under varying temperature and the applied magnetic field, but a fixed microwave frequency (~9 GHz). (b) The structure of bovine complex I (brown: core subunits and grey: supernumerary subunits) is depicted with the iron-sulphur clusters labelled according to their signals (yellow-orange spheres). This figure is taken from Wright *et al*., (Wright et al., 2016).

3.10 Protein analyses of purified mouse complex I

Complex I was previously assessed for its catalytic properties relating to the core subunits (i.e. kinetics, flavin quantification and EPR spectroscopy), which only rely on 14 out of 44 unique subunits. Therefore, in order to analyse the protein content of the purified sample and determine whether the protein composition is conserved from the bovine enzyme, mass spectrometry was employed.

First**,** purified mouse complex I was resolved by SDS-PAGE and visualised by staining the gel with Coomassie Blue. The gel was excised into equal-size bands along the gel and the proteins were identified by peptide mass fingerprinting after tryptic digestion and analysis using a MALDI TOF/TOF mass spectrometer (Figure 3.6 and Table 3.5). This technique has the advantage of identifying a protein band directly and matching its relative position to the molecular size markers (Figure 3.6), ensuring that the peptides observed do not originate from a proteolysis product. 33/44 unique subunits of complex I were identified using this method, of which 9 are core and 24 are supernumerary subunits. Of those core subunits detected, 2 are mitochondrial-encoded: ND5 and ND4 both show relatively low sequence coverage and low number of peptides detected. This is owing to the lack of suitable peptides generated by tryptic digestion due to the hydrophobicity of those proteins. Also detected in the preparation by the MALDI TOF/TOF analyses are proteins of other function that are not considered complex I subunits (Table 3.6).

The subunit composition of the mouse complex I preparation was also examined by LC-MS. In brief, a purified complex I sample was digested with trypsin in solution, then peptides were separated on a reverse-phase HPLC column and introduced directly into the MS using electrospray ionisation. Thereafter, their masses were measured, followed by peptide dissociation and measurement of peptide fragments (Table 3.5). After LC-MS analyses, 39 complex I subunits were detected, of which 10 are core and 29 are supernumerary subunits. As with the MALDI TOF/TOF analysis, only 3 mitochondrial-encoded subunits were detected (ND5, ND4 and ND1), again owing to the poor peptides generated by tryptic digestion.

Complex I subunits have various post-translational modifications, ranging from the removal of N-terminal import sequences to N-terminal acetylation (Carroll et al., 2009). Therefore, in order to observe the post-translational modifications present in mouse complex I, and to see if

modifications are conserved when compared to bovine, subunits were first separated by reverse-phase HPLC (Figure 3.7) and the molecular masses of the individual subunits were measured by ESI-MS. The measured masses were tentatively assigned to subunits of complex I, with the aid of previous work on the bovine enzyme (Carroll et al., 2009). Masses were assigned to 37 subunits (Table 3.7). Comparison of measured and predicted mass data suggest that of the 37 subunits identified, 11 are N-terminal acetylated, 3 are N-terminal formylated, 14 have the import sequence removed, 1 is hydroxylated, 13 have the initiator methionine removed, 1 is dimethylated, 2 have acyl lipid modification modifications and 1 is unmodified. Some subunits contain more than one modification (Table 3.7). Four subunits have been more cautiously assigned to measured masses, as there is a considerable deviation from the expected mass: NDUFV1, NDUFS2, ND4 and ND2. This may be due to a poorer quality signal, sequence discrepancy to the sequence in the database or different modifications than those observed in the bovine enzyme. Nevertheless, their assignment was aided by comparison of the reverse-phase HPLC elution profile of the subunits with data from that of the bovine enzyme (Carroll et al., 2009).

In total, using mass spectrometry, 42/44 unique subunits were identified; all of the supernumerary subunits and most of the core subunits (with the exception of ND3 and ND6, again, due to the poor peptides generated by these proteins). Owing to the full catalytically competency of the enzyme, it can be in no doubt that the full complement of subunits are present. Although contaminating proteins were detected in the preparation, size-exclusion chromatography by way of a monodispersed peak, and BN-PAGE analyses showed that by far, the major component in this in preparation is complex I. In conclusion, this preparation is relatively pure, with most of the subunits identified and all of them inferred, and it is a suitable model for structural and other works.

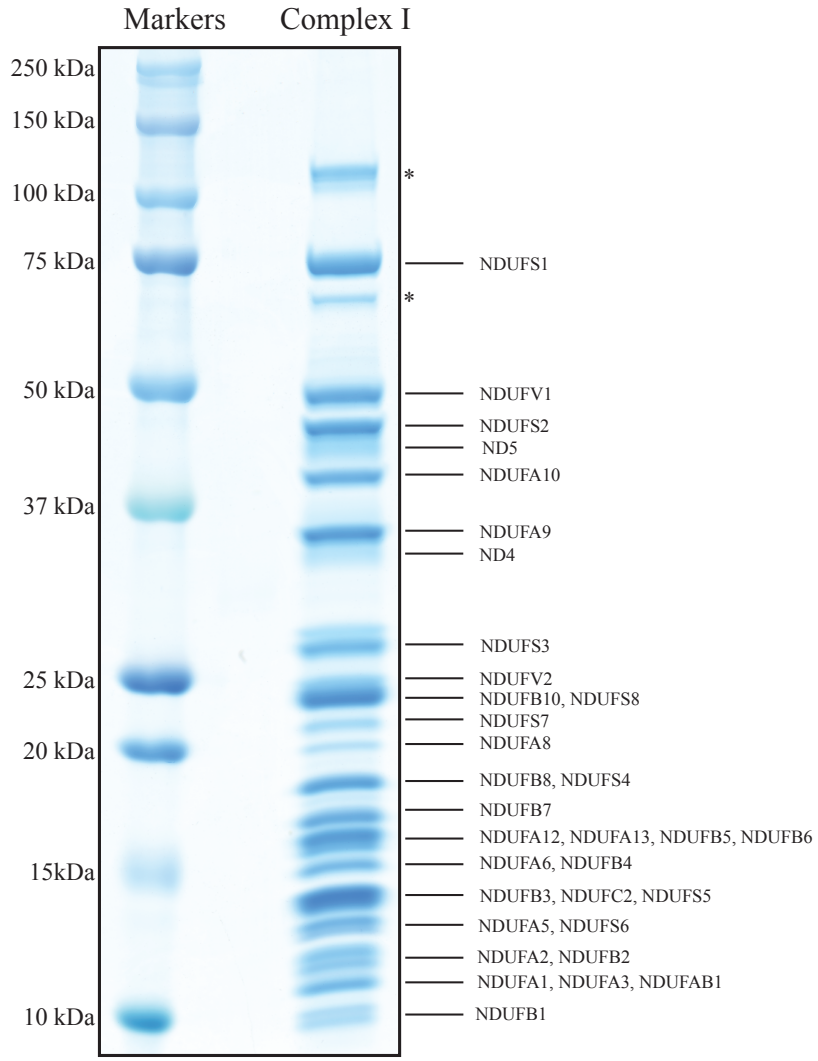


Figure 3.6: SDS-PAGE of purified mouse complex I. Complex I subunits (9 µg) were separated by SDS-PAGE and then were stained with Coomassie Blue. Subunits were identified by peptide mass fingerprinting using a MALDI-TOF/TOF mass spectrometer. * Indicates proteins not considered to be associated with complex I. Known molecular size markers were loaded as a size reference.

Subunit	Number of residues	MALDI TOF/TOF MS/MS			LC MS		
		Number of peptides	Sequence coverage	Score	Number of peptides	Sequence coverage	Score
NDUFA1	70	1	14%	66/34	7	42%	227
NDUFA2	98	5	38%	191/33	12	78%	538
NDUFA3	83	2	21%	74/34	5	61%	105
NDUFA5	115	3	45%	209/33	13	86%	650
NDUFA6	130	4	33%	203/32	24	96%	1255
NDUFA7	112	-	-	-	32	100%	1880
NDUFA8	171	5	39%	275/33	13	52%	553
NDUFA9	342	12	41%	836/34	28	72%	1988
NDUFA10	320	11	44%	723/33	21	45%	1343
NDUFA11	140	-	-	-	2	10%	236
NDUFA12	145	2	6%	59/34	18	86%	1069
NDUFA13	143	3	22%	138/34	22	85%	2347
NDUFAB1	88	4	37%	216/34	1	20%	71
NDUFB1	56	2	32%	85/34	5	51%	130
NDUFB2	72	2	12%	52/33	1	12%	80
NDUFB3	103	2	14%	51/34	9	37%	171
NDUFB4	128	6	46%	298/34	18	70%	1053
NDUFB5	143	3	37%	114/34	10	50%	568
NDUFB6	127	5	40%	175/34	17	76%	773
NDUFB7	136	2	10%	134/30	7	51%	787
NDUFB8	157	7	58%	449/33	11	57%	867
NDUFB9	178	-	-	-	20	68%	716
NDUFB10	175	4	26%	235/33	14	60%	756
NDUFB11	122	-	-	-	15	76%	690
NDUFC1	49	-	-	-	-	-	-
NDUFC2	120	4	27%	166/34	12	60%	694
NDUFS1	704	12	21%	653/33	54	51%	4537
NDUFS2	430	10	33%	729/33	24	53%	1564
NDUFS3	228	12	57%	818/33	22	55%	1448
NDUFS4	133	6	39%	449/33	13	49%	506
NDUFS5	105	1	9%	85/34	11	49%	309

NDUFS6	96	4	39%	278/33	8	73%	827
NDUFS7	189	5	17%	217/33	12	38%	733
NDUFS8	178	4	29%	318/33	9	44%	350
NDUFV1	444	11	28%	574/34	32	47%	1522
NDUFV2	217	5	29%	311/34	13	51%	1504
NDUFV3	69	-	-	-	8	71%	222
ND1	318	-	-	-	3	8%	140
ND2	345	-	-	-	-	-	-
ND3	115	-	-	-	-	-	-
ND4	459	3	7%	149/33	3	7%	118
ND4L	98	-	-	-	-	-	-
ND5	607	3	7%	84/32	5	8%	253
ND6	172	-	-	-	-	-	-

Table 3.5: Peptide-based protein identification. The number of peptides detected from each subunit is given together with the sequence coverage of the mature protein (number of residues column). For the two analysis methods, the score given for each of the subunits is the sum of the peptide score with denominator (denominator value for LC-MS score is 21) representing the 95% confidence threshold ($p < 0.05$).

Protein	Number of residues	MALDI TOF/TOF MS/MS		
		Number of peptides	Sequence coverage	Score
SLC4A1	929	4	5%	180/33
ACADVL	656	12	25%	796/33
ATP5A1	503	6	16%	172/33
DLST	454	2	5%	63/33
COX2	227	1	4%	52/34
PLCG1	132	2	7%	63/32
COX4I1	169	4	25%	158/34
COX7A1	80	1	16%	91/34
COX7C	63	1	14%	44/34
NDUFAF2	168	4	23%	80/33

Table 3.6: Identification of contaminants in complex I preparation. The number of peptides detected from each subunit is given together with the sequence coverage of the protein. The score given for each of the subunits is the sum of the peptide score with denominator representing the 95% confidence threshold ($p < 0.05$).

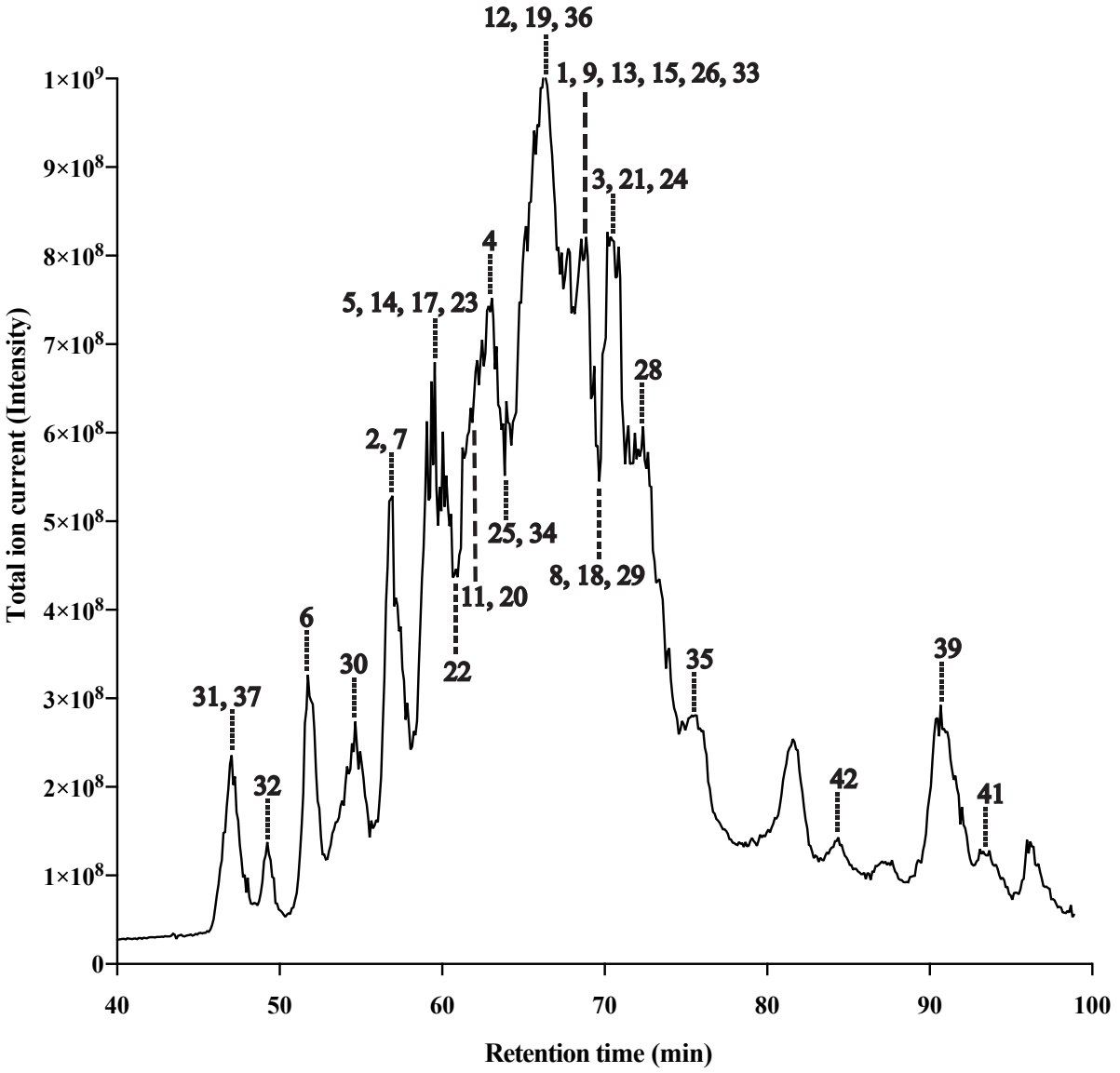


Figure 3.7: Analysis of mouse mitochondrial complex I by LC-MS. Complex I (~100 µg**)** subunits were separated by reverse-phase HPLC and detected by ESI-MS. The ion intensities of the eluents are indicated, and complex I subunits are identified and numbered according to Table 3.7.

Subunit	Mass		Δ mass discrepancy	Modifications
	Calculated	Observed		
1. NDUFA1	8138.5	8138.9	0.4	-
2. NDUFA2	10826.5	10826.2	-0.3	-Met, +N-acetyl
3. NDUFA3	9241.7	9242.2	0.5	-Met, +N-acetyl
4. NDUFA5	13270.5	13270.9	0.4	-Met, +N-acetyl
5. NDUFA6	15193.6	15193.8	0.2	-Met, +N-acetyl
6. NDUFA7	12486.4	12487.1	0.7	-Met, +N-acetyl
7. NDUFA8	19852.9	19852.9	0.0	-Met, +4×Cys-Cys
8. NDUFA9	38825.8	38829.1	3.3	-ΔImport (1-35)
9. NDUFA10	36943.0	36947.7	4.7	-ΔImport (1-35)
10. NDUFA11	14851.1	-	-	-
11. NDUFA12	17128.4	17129.2	0.8	+N-acetyl
12. NDUFA13	16770.4	16770.9	0.5	-Met, +N-acetyl
13. NDUFAB1	10662.1	10661.3	-0.8	-ΔImport (1-68), +4PP-acyl chain
14. NDUFB1	6822.9	6823.3	0.4	-Met
15. NDUFB2	8491.4	8491.4	0.0	-ΔImport (1-33)
16. NDUFB3	11561.0	-	-	-
17. NDUFB4	14992.1	14992.4	0.3	-Met, +N-acetyl
18. NDUFB5	16858.4	16858.6	0.2	-ΔImport (1-46)
19. NDUFB6	15425.9	15426.1	0.2	-Met, +N-acetyl
20. NDUFB7	16405.8	16406.4	0.6	-Met, +N-myristoyl, +2×Cys-Cys
21. NDUFB8	18745.1	18746.8	1.7	-ΔImport (1-29)
22. NDUFB9	21894.9	21892.5	-2.4	-Met, +N-acetyl
23. NDUFB10	20888.6	20889.7	1.1	-Met, +2×Cys-Cys
24. NDUFB11	14343.0	14343.9	0.9	-ΔImport (1-29)
25. NDUFC1	5752.6	5753.1	0.5	-ΔImport (1-27)
26. NDUFC2	14205.7	14206.8	1.1	+N-acetyl
27. NDUFS1	77210.6	-	-	-
28. NDUFS2	49257.6	49278.9	21.3	ΔImport (1-27), +dimethyl-R85
29. NDUFS3	26479.0	26479.9	0.9	-ΔImport (1-35)
30. NDUFS4	15325.3	15325.4	0.1	-ΔImport (1-42)

31. NDUFS5	12512.5	12514.4	1.9	-Met, +2×Cys-Cys
32. NDUFS6	10777.0	10775.8	-1.2	-ΔImport (1-20)
33. NDUFS7	21032.5	21033.6	1.1	-ΔImport (1-35), +Hydroxyl-R87
34. NDUFS8	20442.2	20439.5	-2.7	-ΔImport (1-35)
35. NDUFV1	48626.5	48647.7	21.2	-ΔImport (1-20)
36. NDUFV2	23847.4	23846.8	-0.6	-ΔImport (1-31)
37. NDUFV3	8027.8	8029.0	1.2	-ΔImport (1-35)
38. ND1	36087.3	-	-	-
39. ND2	38780.4	38794.7	14.3	+N-formyl
40. ND3	13246.9	-	-	-
41. ND4	51909.8	51924.5	14.7	+N-formyl
42. ND4L	10634.7	10633.6	-1.1	+N-formyl
43. ND5	68502.8	-	-	-
44. ND6	18654.2	-	-	-

Table 3.7: Molecular mass measurement of complex I subunits. Measured masses were assigned to complex I subunits according to the mass discrepancy between the observed masses and calculated masses (mature protein with known modifications), as shown in previous work (Carroll et al., 2009). Modification of the observed masses includes (-) cleavage of import sequences or residues, or (+) or addition of post-translation modifications. Disulphide bridges are presented as Cys-Cys.

3.11 Discussion

3.11.1 Comparison to other preparations of mammalian complex I

Although the mouse enzyme displayed a higher catalytic rate compared to the ovine enzyme (Letts et al., 2016b), it was substantially less active than the current preparation of the bovine enzyme (Blaza et al., 2018). The reason for this difference in activity is unclear at present. Just as the bovine complex I preparation was improved over time, the current mouse complex I preparation could be improved, by changing various parameters, such as pH, lipid composition and buffer used.

Molecular mass measurements of individual subunits of mouse complex I reveal post-translational modifications, with many examples confirmed by peptide evidence in LC-MS that were identical to that of the bovine and rat enzymes (Bridges et al., 2017; Carroll et al., 2009). In agreement with previous work in rat (Bridges et al., 2017), peptide-based identification showed that no evidence for alternative isoforms of mouse heart complex I subunits were present, and NDUFV3 was only present it is canonical 10 kDa form.

Another complication in this preparation is the presence of trace contaminants that may hinder accurate kinetic characterisation. The preparation contained components of complex IV and complex V (Table 3.6). These contaminant proteins constitute minor components in the preparation as shown by BN-PAGE (Figure 3.3).

Complex I from ovine was explored as an alternative to the bovine model for structural studies by another group (Letts et al., 2016b). Their preparation resulted in a complex I, of comparable purity to the mouse enzyme. However, the catalytic activity was lower than both the bovine and mouse complexes I. Moreover, the detergent (LMNG) explored in the solubilisation of ovine membranes does not result in mostly monomeric complex I in mouse (Figure 3.1), thus affecting the yield of complex I obtained.

Complex I has also been purified from the genetically amenable rat model (Bridges et al., 2017). Unlike mouse complex I, however, the rat enzyme is unstable in DDM and therefore was purified in LMNG. Furthermore, due to ATPase contamination in the preparation, an additional affinity chromatography (IF_1 present in the matrix of the column) step was included

(compared to the mouse purification) to remove the contaminant. Finally, the catalytic activity of the rat enzyme was significantly less than its mouse counterpart (Mohammed, 2015).

3.11.2 Conclusion

In this latest preparation, mouse complex I was purified similarly to the bovine enzyme, but scaled-down significantly. In summary, membranes were solubilised in DDM, followed by resolving the complexes in an anion exchange column (supplemented with phospholipids) and finally, complex I was separated from complex IV in a size-exclusion column. The resulting mouse complex I was pure and highly active, with characteristic spectroscopic features indistinguishable from those of the well-characterised bovine enzyme. In conclusion, mouse complex I prepared according to this method provides a promising new model for future structural and mechanistic studies to understand complex I function, and dysfunction (through the use of disease model) in relation to human pathologies.

3.11.3 Future prospects

The main advantage of working with mouse as a model for complex I studies rather than other mammalian systems, is that genetic manipulation can be easily carried on mouse, trapping the enzyme in catalytically-relevant or assembly-stalled states (see table 3.1).

During the protein content analyses of the preparation, it became clear that NDUFS6 subunit of mouse complex I possess a tri-histidine motif at its N-terminus, which is not present in the bovine enzyme (Figure 3.8). Therefore it may be possible to use Ni-exchange chromatography to purify the complex I, as already done in complex I of *Pichia pastoris*, which also has a naturally occurring tri-histidine tag. This may prove crucial in the future for purifying complex I from assembly intermediates that have altered behaviours on anion-exchange chromatography or to improve purity. This may be particularly relevant, as discussed in chapter 6, for the problematic purification of complex I from the NDUFS4 subunit deletion mouse strain (Kruse et al., 2008). Complex I from this strain is reported to be either destabilised or to form a stalled assembly intermediate (Calvaruso et al., 2012). Therefore, it would be ideal to have an alternative purification method, in case the one discussed in this chapter is not suited for complex I mutants.

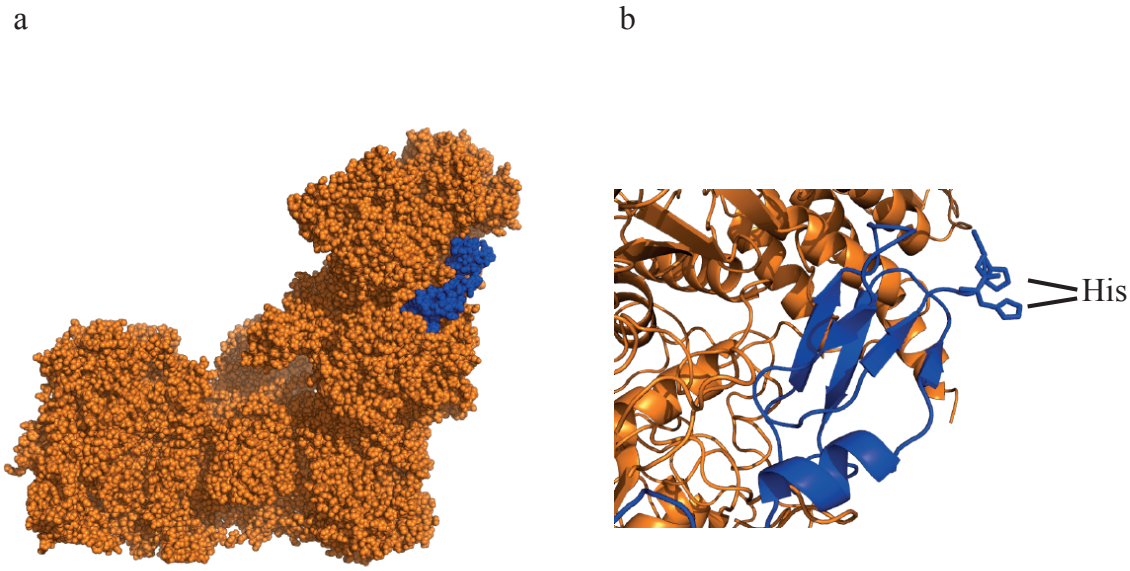


Figure 3.8: Representation of mouse complex I with highlighted NDUFS6. The structure is discussed in detail in chapter 4 (a) Complex I presented as spheres with NDUFS6 as blue and the rest of the enzyme as orange. NDUFS6 is in cartoon representation showing the position of two modelled solvent-exposed histidines (indicated). The modelled structure depicted is of the active mouse complex I in chapter 4.

Chapter 4: Structures of mouse complex I by cryo-EM

4.1 A brief Introduction to cryo-EM

Cryo-electron microscopy (cryo-EM) is a technique used to image thin frozen specimens using electrons from a transmission electron microscope. Recent developments in cryo-EM (e.g. new detectors and better software algorithms for 3D reconstruction of cryo-EM maps) has resulted in an exponential increase in the number of new structures available in the protein data bank. As of the 6th of March 2018, there have been a total of 1991 structures released using this method (Figure 4.1).

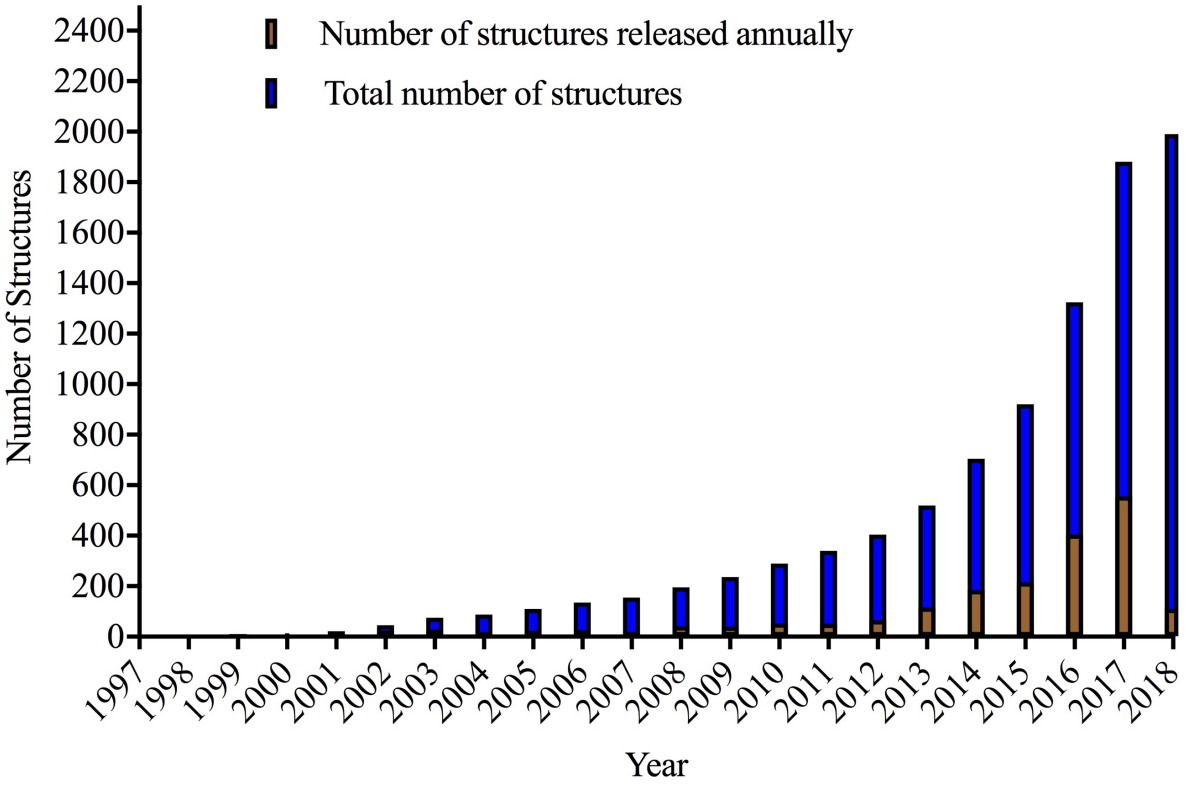


Figure 4.1: Number of structures released that were determined using cryo-EM. The values are taken from the protein databank (Berman et al., 2000).

Cryo-EM has facilitated the determination of new high-resolution structures, which were previously only attainable by X-ray crystallography and NMR, such as the 2.2 Å resolution structure of β-galactosidase (Bartesaghi et al., 2015), or the structure of glutamate

dehydrogenase at 1.8 Å resolution (Merk et al., 2016). Although X-ray crystallography remains the most widely used technique for high resolution structural determination, it is limited by how readily a protein crystallises. Countless parameters have to be adjusted and optimised to attain protein crystals (e.g. salt concentrations, precipitants and other additives), and as a consequence, the structure observed may diverge from the native structure. Attempts to crystallise mammalian complex I in this laboratory failed for many years, with a membrane subdomain of bovine complex I at only 6.8 Å resolution being the sole structure solved by this method (Zhu et al., 2015).

The key advantage that has allowed single particle cryo-EM to propagate as a common technique to obtain near-atomic resolution structures is its simple work flow (Figure 4.2). A small volume of purified protein is added to a grid and then blotted to create a thin film of liquid that is then plunge-frozen in liquid ethane, leaving the protein embedded in vitreous ice. The frozen samples are then imaged with a 300 keV microscope producing 2D experimental projections on a direct electron detector. The resulting experimental projections are then processed computationally to construct 3D density maps. Finally, a coordinate model is built with the generated map.

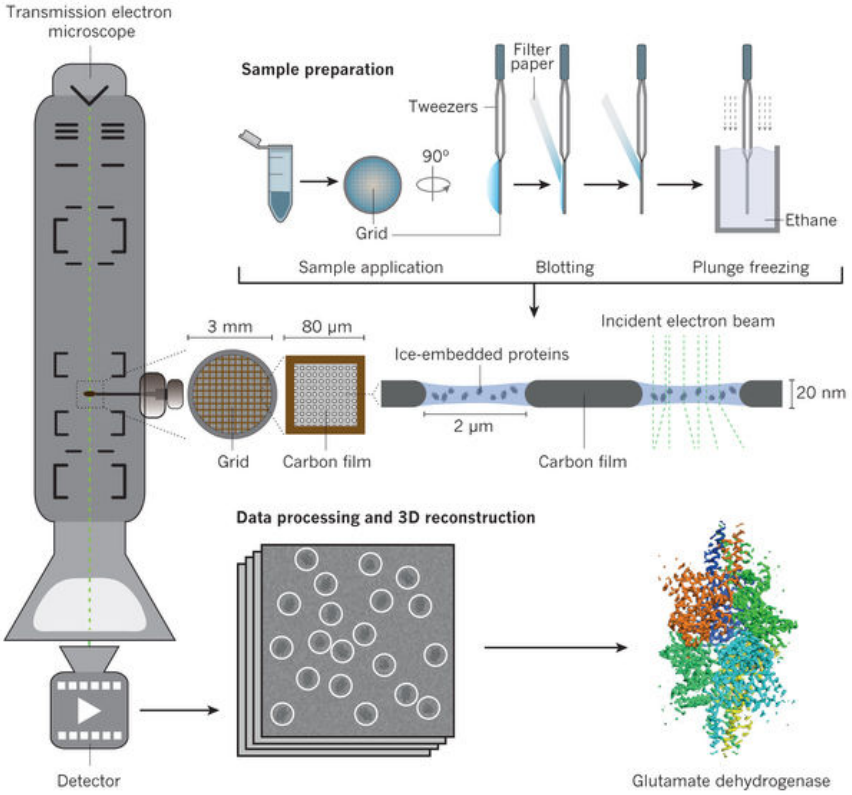


Figure 4.2: A schematic workflow for a typical cryo-EM project. This figure was taken from Fernandez-Leiro and Scheres (Fernandez-Leiro and Scheres, 2016).

The first major advance in single particle cryo-EM involved the preservation of biological samples. Prior to cryo-EM, biological samples were first coated with heavy salts such as uranyl acetate. This coat increased the contrast of the images of the specimen against the background by increasing its scattering potential. However, all internal features of the samples were lost because only the heavy salt stain envelopes were being imaged. Also, protein samples are hydrated in their native states, therefore removing water by this staining method may lead to changes in structure that do not reflect the native state (Thompson et al., 2016).

Dubochet and co-workers devised a protocol for rapidly freezing biological samples and thus avoiding the formation of ice crystals, which are damaging to cells and their proteins (Adrian et al., 1984). The process first developed, on viruses, was termed vitrification and it uses liquid ethane (liquefied by liquid nitrogen), which has a greater freezing efficiency than liquid nitrogen to freeze samples. The method works because the rate of sample freezing exceeds the rate at which the hydrogen bonding networks formed by water molecules during ice formation can rearrange. Thus, the molecules are trapped in an amorphous or vitreous state of non-crystalline ice. In the vitreous ice a) the proteins are in a hydrated state, b) the internal structures

in the protein become visible (unlike in negative stain) and c) the cryogenic temperatures used mitigate radiation damage by incident electrons (Henderson, 1990).

Another major recent advance was the development of new direct electron detectors that increased the detection efficiency of the detectors. Previously, either film or charge coupled device (CCD) camera were used for detection of electrons. The main disadvantage of the CCD camera was that electrons could not be detected directly, but the signal had to be first converted to light by a scintillator, then transferred to the CCD camera to digitise and transfer the signal (Figure 4.3). High energy electrons are first converted to photons because they are damaging to the detector, and as a consequence of this, information is lost first in the conversion and second by deviation in the pathways of the photons (McMullan et al., 2016), causing a spread in the image information (blurring). Alternatively, electrons can be detected directly with film. However, the film has to be processed after data collection, therefore settings such as exposure and astigmatism, which are critical for obtaining high-resolution structures, cannot be optimised during the imaging session.

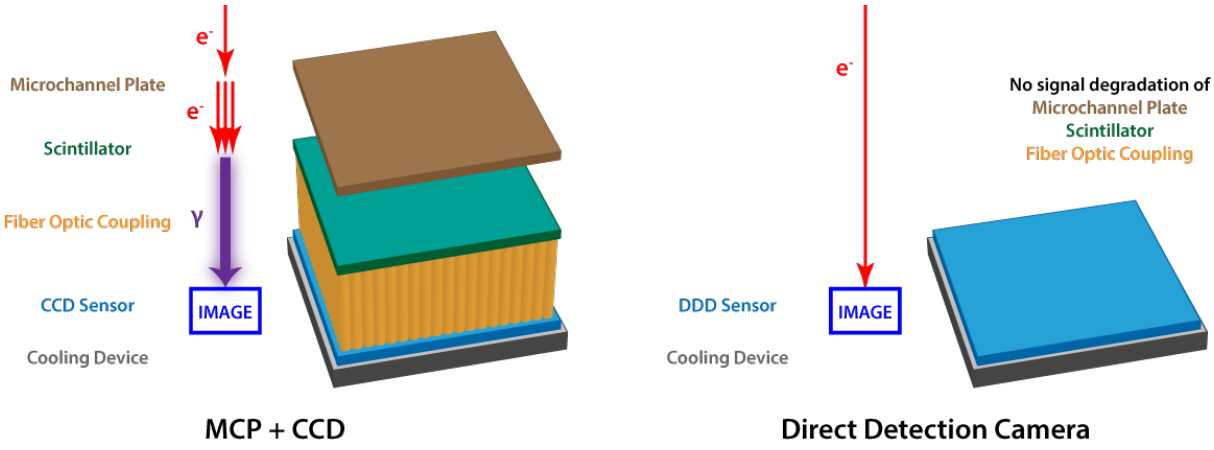


Figure 4.3: A schematic of how electrons are detected by CCD sensors and direct electron detectors. This figure was taken from Direct Electron (www.directelectron.com/products/lv-series).

The direct electron detector detects electrons directly using active pixel in the immediate vicinity of the electron pathway, thus avoiding information loss from converting the electron signal into photons. The exact mechanism of how direct electron detectors function is described in McMullan *et al.,* (McMullan et al., 2016). The technological advancement in these direct

detectors is reflected by their improved quantum efficiency over film and CCD cameras (McMullan et al., 2014). In addition, because of the fast readout of the electron signal by the direct detector, successive frames for each exposure or movie can be used to correct for beam-induced motion and radiation damage (Glaeser, 2016).

Another major advance in electron microscopy was the development of reconstruction algorithms that allowed for the generation of 3D cryo-EM maps from 2D experimental projections (Orlova and Saibil, 2011). In order to do this, a) protein 'particles' must be randomly orientated in vitreous ice to allow 3D information to be gained from the experimental projections and b) the relative orientations of these projections must be known.

The main method used for cryo-EM reconstruction is called projection matching (Figure 4.4). The process requires an initial reference map that is used to generate model projections along the structure, and these are then used as references against which experimental projections are compared (Orlova and Saibil, 2011). Once the experimental projection agrees with the model projections, then relative orientations can be assigned to the experimental projections, which are then back-projected into a 3D volume, thus generating a new map. The new map is then used iteratively for projection matching, refining the model until there are no further improvements in the accuracy of the orientational assignments.

Current programs such as RELION do not work by matching one experimental projection to one reference projection, but rather use a statistical method to assign weighted probabilities to all the reference projections for a given experimental projection (Scheres, 2016). With each successive iteration, the weighted probability values change such that an experimental projection overwhelmingly matches a single reference projection. This method is useful because the experimental projections are inherently noisy (i.e. low signal to noise) and therefore a single projection may be assigned incorrectly to a model projection. Furthermore, RELION guides the reconstruction by calculating a probability-based score for the different possible orientations of an experimental projection based on prior knowledge (i.e. signal and noise follow a normal distribution) (Scheres, 2016).

A minor point is that the reconstruction in RELION is performed in Fourier space (amplitude vs. spatial frequency). A real space 2D projection is a 2D slice (perpendicular to the projection view) of a 3D transform of the object (Scheres, 2016). This is because it is computationally less

expensive to work in Fourier space and reverse transform the volume into a real space map than it is to work in real space, for reasons which are beyond the scope of this thesis. The results are the same nevertheless.

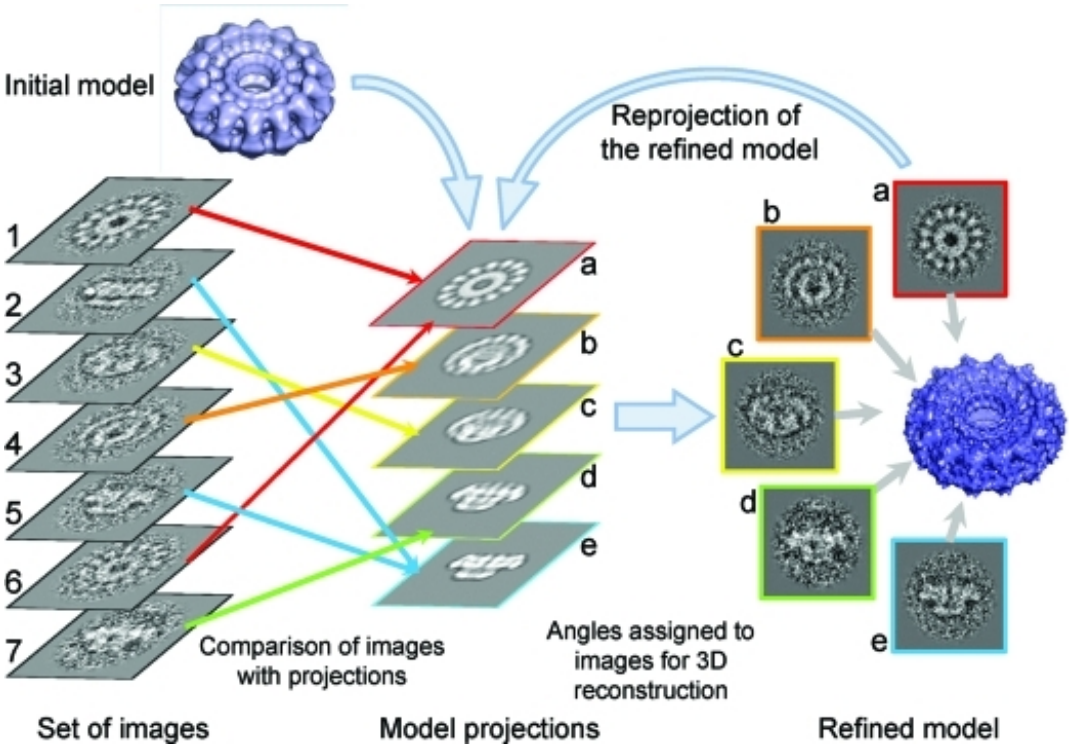


Figure 4.4: Projection matching. An initial model is used to make 2D reference projections, which are then matched to experimental projections. With defined orientations, the experimental projections are then used to make a new 3D map. In subsequent refinement steps, the new map is then iteratively re-projected and compared to the experimental projections to ascertain more accurate orientation assignment and thus a better map. This figure was taken from Orlova and Saibil (Orlova and Saibil, 2011).

4.2 Aims of this chapter

The work in this chapter aimed to:

- Purify mouse complex I in two different states, ‘active’ and ‘de-active’ for cryo-EM data collection.
- Generate high-resolution maps from cryo-EM data using the current data processing software (RELION) for model building.
- Analyse and compare the structures in the two defined states, in order to identify structurally, functionally and catalytically-relevant information.

4.3 Image processing

Active and de-active complex I (4-5 mg mL^{-1}) were purified from mouse heart (as described in Chapter 2.7). The proteins were applied to onto PEGylated gold grids and then plunge frozen in liquid ethane (Blaza et al., 2018; Meyerson et al., 2015; Russo and Passmore, 2014). The grids were imaged using a Titan Krios microscope with a K2 detector at a calibrated pixel size of 1.05 Å and 1.33 Å for the active and de-active mouse complex I, respectively (described in Chapter 2.13).

Unless stated otherwise, the image processing steps described are for the data of the active mouse complex I.

The electron beam from the microscope causes in-plane motion and radiation damage to the sample, which ultimately affects the retention of high-resolution information in the micrographs (Glaeser, 2016). Therefore, in order to account and correct for this, MotionCor2 was used on whole micrographs (Zheng et al., 2017). Following motion correction, the contrast transfer function (CTF) of the micrographs were estimated using Gctf (Zhang, 2016), in order to correct for aberrations in the microscope that results in the degradation of information (i.e. blurring of the images on the micrographs). Protein 'particle' coordinates were automatically selected using ten bovine complex I 2D class averages as (kindly provided by Dr. Hannah Bridges), which were low passed filtered to 20 Å to prevent picking bias, and then manually curated (Figure 4.5). As seen in the representative micrograph, the particles are well distributed on the grid and randomly orientated, with no noticeable preferential orientation. This makes it ideal for 3D reconstruction.

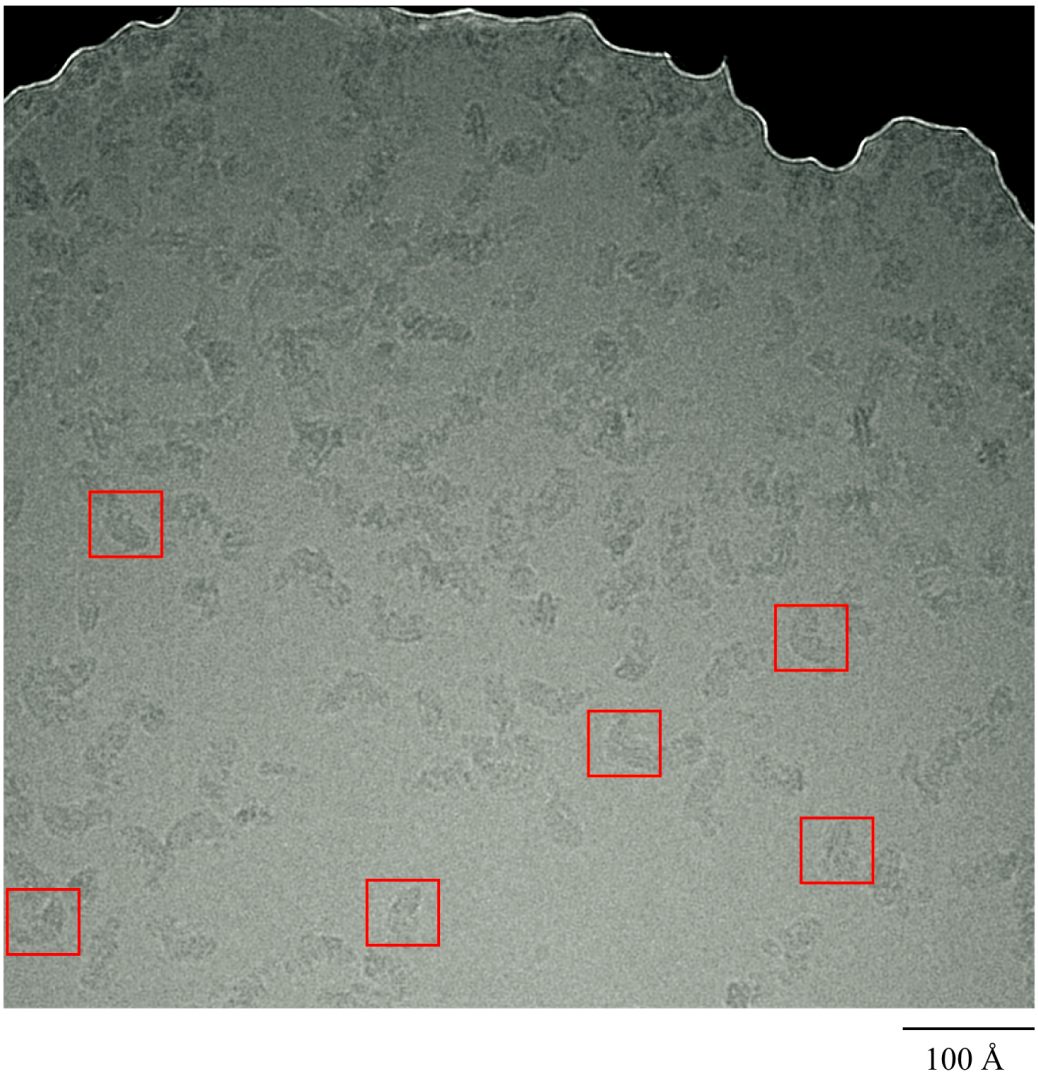


Figure 4.5: A representative micrograph of mouse complex I imaged on a PEGylated gold grid. A few examples of complex I particles to be selected are shown in red. A scale bar is shown below for reference. The image was low passed filtered to 20 Å in RELION to increase the contrast between the protein and the solvent (Kimanius et al., 2016).

Once particle coordinates had been selected, the particles were extracted and normalised, with the contrast of the particles inverted (background is black and protein particles are white). The extracted particles were then subjected to several rounds of reference-free 2D classification. This procedure requires that the number of classes 'clusters' to be set (e.g. 100) and that those particles that have similar in-plane orientations are averaged together, enhancing the signal to noise of the particles (Frank, 1975) (Figure 4.6). This procedure is useful in removing non-complex I particles from the dataset, as they do not form large clusters and therefore have low single to noise features.

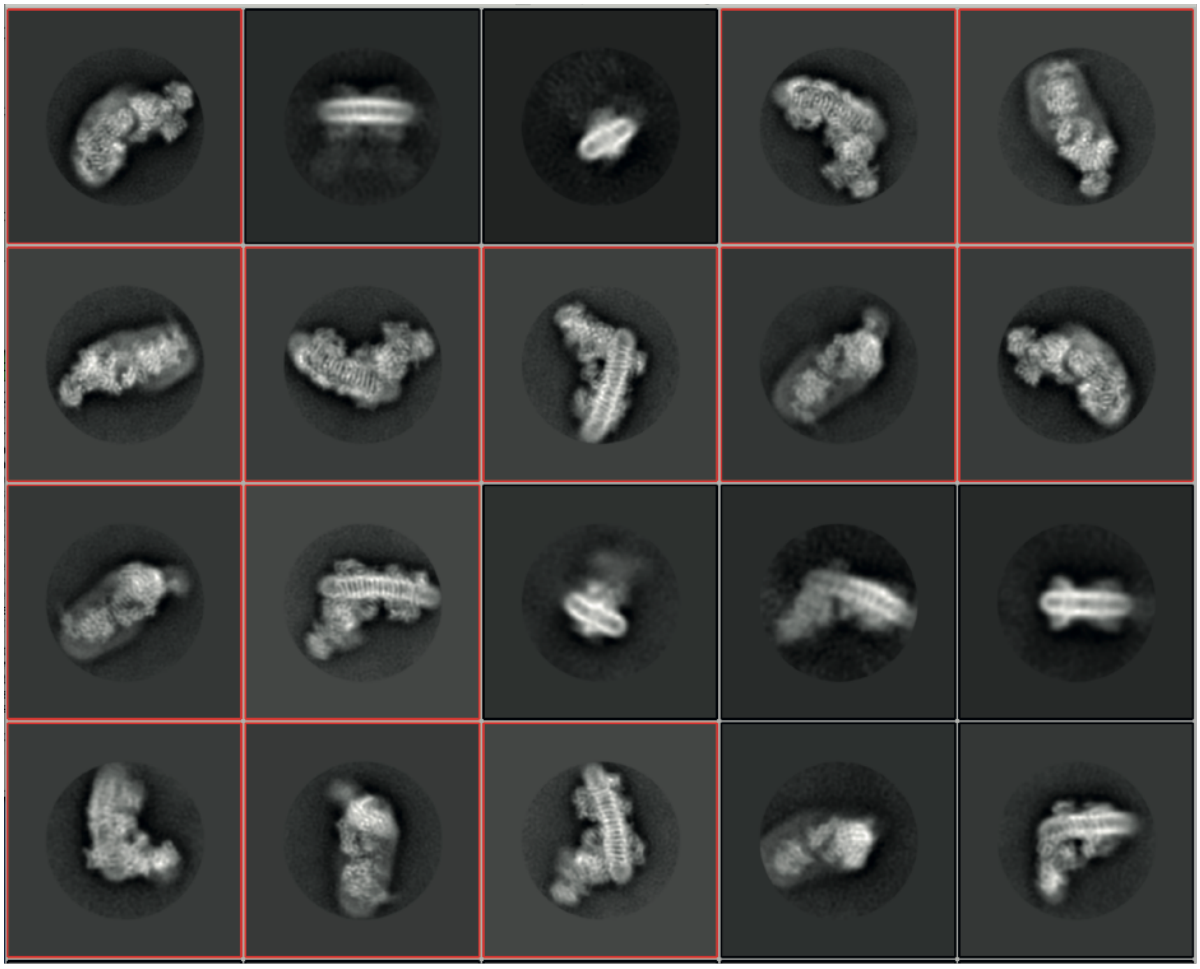


Figure 4.6: An example of a set of 2D class averages. Good classes are highlighted by the red box whilst bad/ambiguous classes are not. The 2D class averages were generated in RELION (Kimanius et al., 2016).

First, averaged particles that resembled complex I were selected over those that did not. Second, 'Good classes' of complex I particles were easily distinguished from 'bad classes' by the presence of secondary structure elements in the 2D class averages (Figure 4.6). In good classes, transmembrane helices are clearly visible, whilst they are not visualised in bad classes. Furthermore, the contrast between the averaged particles and the solvent is enhanced in good classes, because the classes are more homogenous and therefore the signal relative to the disordered solvent is improved. In contrast, bad classes are blurred in the images, which is an indication of either a lack of homogeneity or broken protein (Scheres, 2016). Classes were selected for further refinement based on these criteria.

Following movie refinement, which tracks the movements of selected particles, the RELION 'polishing' procedure, was applied using default parameters for B-factor weighting to correct for per particle beam-induced motion, and radiation damage by appropriately down-weighting the contribution of later frames (Kimanius et al., 2016). In addition, only frames 1 –12 out of 25 were considered for further refinement (total dose 24 electrons $Å^{-2}$), as they are less radiation damaged than later frames.

In order to obtain a homogenous complex I class The particles were then subjected to 3D classification (with an angular sampling to 3.7°), with 5 classes provided (Figure 4.7). The two major classes (containing 70% of the particles,) were indistinguishable by using the 'Fit in Map' function in UCSF Chimera (Pettersen et al., 2004b). The other minority classes are low-resolution structures, as evident by their apparent lack of secondary structure features. Therefore, only the two selected classes were combined and taken forward to a final 3D auto-refinement.

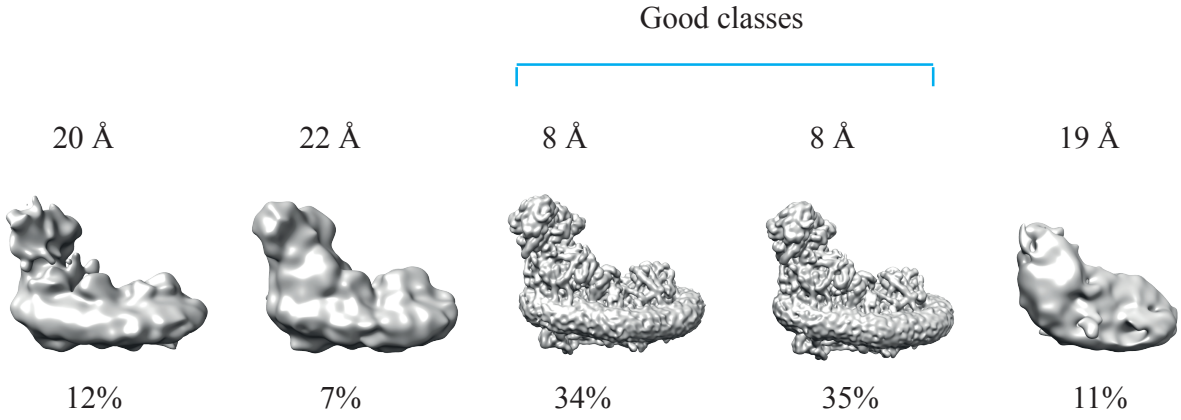


Figure 4.7: 3D classification of particles. 3D classes are shown with their respective class distribution and resolution. The good classes taken further for the final refinement are indicated.

Finally, selected 3D classes were subjected to a final 3D refinement using solvent flattened Fourier shell correlation (FSC) curves with a mask to remove the influence of the solvent and improve the accuracy of orientational assignments. Post-processing was performed on the generated map to yield a resolution of 3.3 Å based on the gold standard FSC = 0.143 criterion (Scheres and Chen, 2012)**.** Below is an overall scheme of the classifications and refinements carried out on both the active and de-active mouse complex I datasets, with the latter carried out by Dr. James Blaza (Figure 4.8).

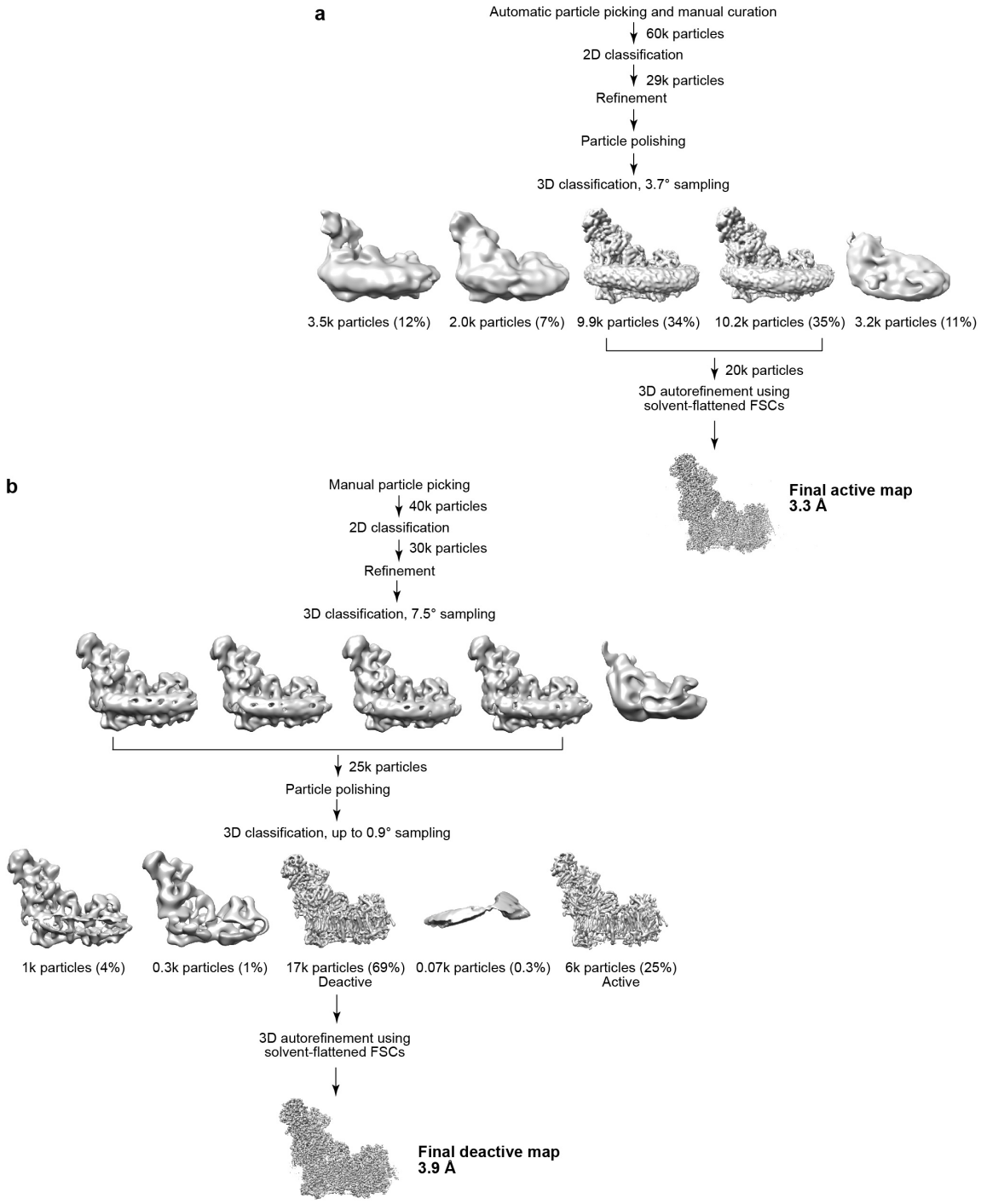


Figure 4.8: Classification and refinement scheme for the active and de-active mouse complex I. All of the above procedures were carried out in RELION (Kimanius et al., 2016).

4.4 Structure of mouse complex I

The determined 3.3 Å resolution of the cryoEM density map for active mouse complex I allowed for a near complete atomic model of the enzyme to be built (Figure 4.9). Most of the model building was carried out by Dr. James Blaza (MRC-MBU, University of Cambridge), based on the previous structure of the bovine enzyme (Blaza et al., 2018). Globally, the subunits organisation of the mouse complex is identical to that of the bovine, ovine, porcine and human complex I structures (Fiedorczuk et al., 2016; Guo et al., 2017; Wu et al., 2016b; Zhu et al., 2016b).

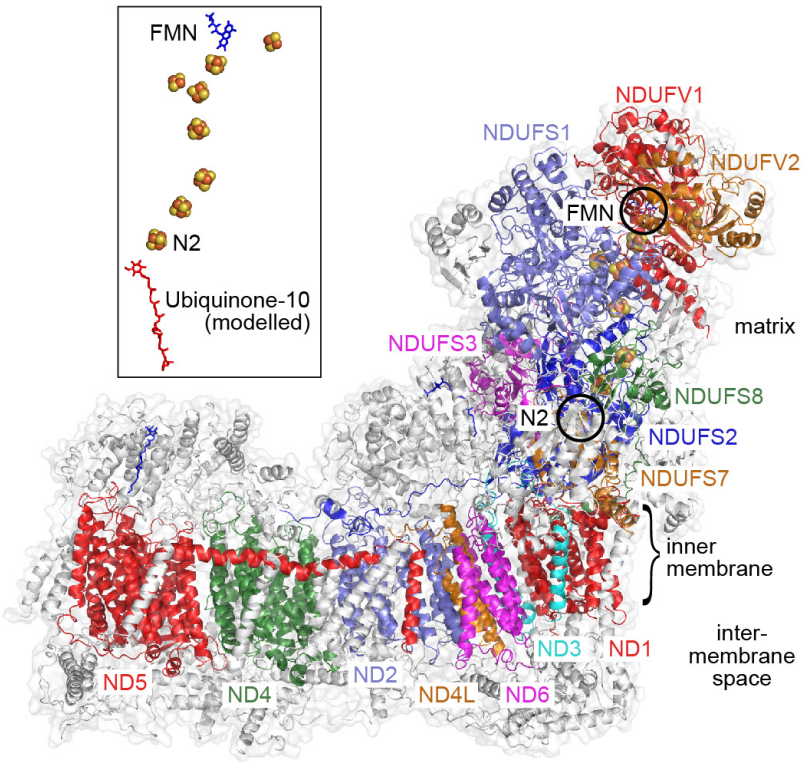


Figure 4.9: Overview of the structure of mouse complex I. The fourteen core subunits are in colour and labelled accordingly and the 26 supernumerary subunits are in white. The positions of the flavin mononucleotide, at the NADH binding site, and FeS cluster N2, the immediate electron donor to ubiquinone, are indicated. The structure of the cofactor chain, together with the modelled structure of ubiquinone-10 (Fedor et al., 2017), is shown separately in the inset also.

The structure contains many well-defined features and side chain densities (Figure 4.10). This enabled confident assignment and modelling of residues with side chains. In total, 96% of the

residues have been modelled in the structure, of which 98% have been modelled in the 14 core subunits and 94% in the 31 supernumerary subunits (Table 4.1). However, there are several stretches of more than 10 residues that are unresolved in the structure, plausibly because of their inherent flexibility, such as the amino and carboxy termini of NDUFS1, NDUFS3, NDUFS7, NDUFA2, NDUFA6, NDUFB3, NDUFB7, NDUFB11 and NDUFV3 (Table 4.1). In addition, 28 residues in the middle of NDUFB6 are poorly resolved and therefore not modelled in. The model also contains known amino-acid modifications, aided by previously measured molecular masses of complex I subunits in mouse (Table 3.7), such as formylation and acetylation of N-terminal residues, disulphide bonds between cysteine residues, dimethyl-Arg85 in NDUFS2, and acyl groups attached to subunits NDUFAB1α and NDUFAB1β (Table 4.1). The supernumerary subunits also contain bound zinc (NDUFS6), and two nucleotides in NDUFA9 and NDUFA10. Finally, there are 20 phospholipids scattered around the structure.

Data for the de-active mouse complex I were processed by Dr. James Blaza. All of the processing steps were performed according to the RELION pipeline procedure (Fernandez-Leiro and Scheres, 2017), as described in Chapter 2.14. The global resolution of the map was estimated to be 3.9 Å according to the gold standard FSC threshold (0.143) (Scheres and Chen, 2012).

The de-active mouse complex I model was produced by Dr. James Blaza, using the active mouse complex I structure as a template (Table 4.1). A comparison of the refinement statistics of active and de-active mouse complex I are presented in Table 4.2. As an additional measure of validation, both the active and de-active mouse complex I models were compared to their respective maps, and there is a clear agreement between the maps and the models (i.e. the map fits into the density well) (Figure 4.11).

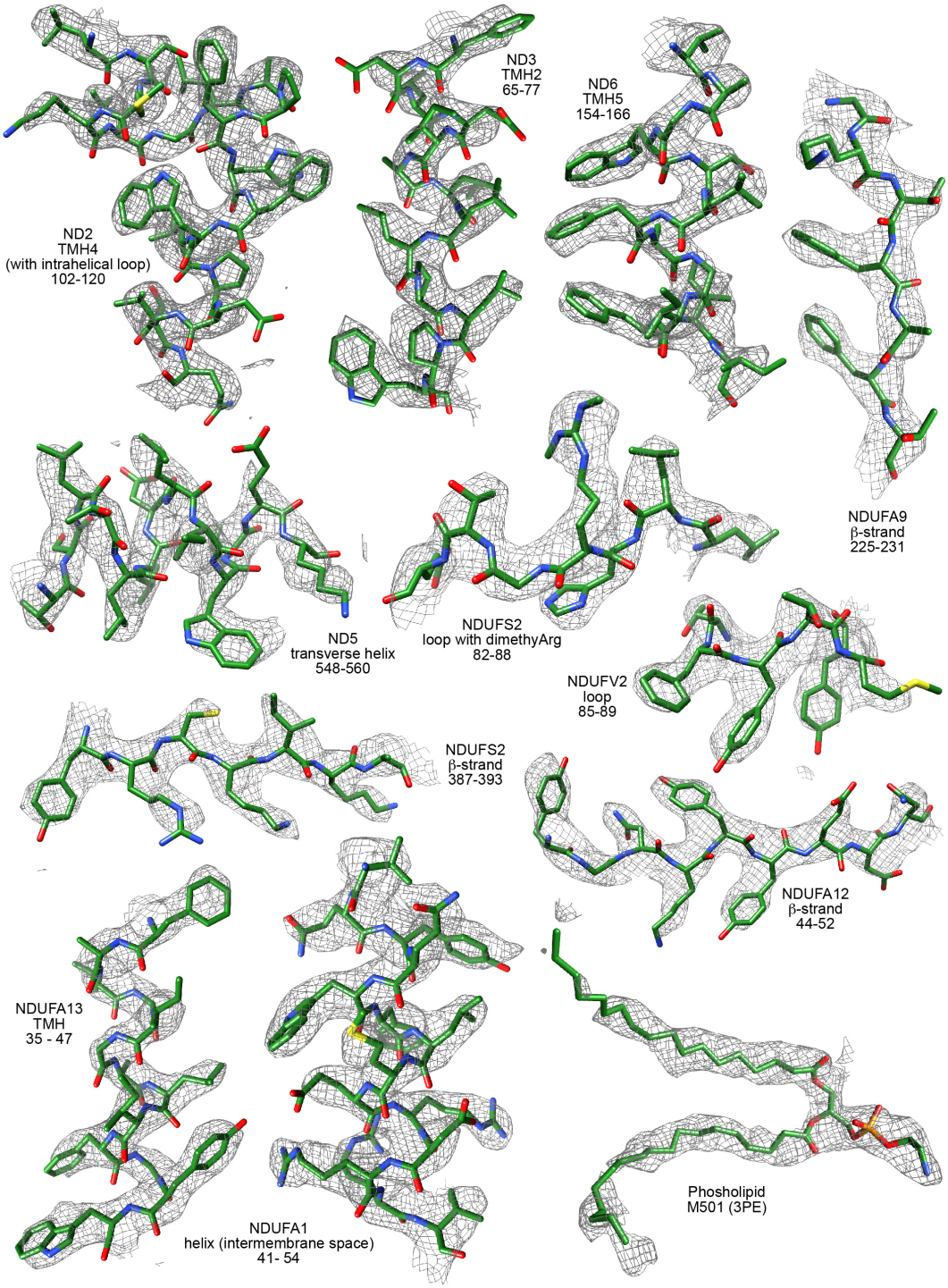


Figure 4.10: Example densities drawn from different parts of the map of active mouse complex I. Densities were carved at 2 Å radius with threshold level of 0.065 and 0.04 for the protein and phospholipid, respectively.

Subunit	Alternative names	Chain	Total residues	Modelled residues	% Modelled	Notes
NDUFV1	51 kDa, Nqo1, NuoF	F	444	10-437	96.4	FMN, 4Fe4S
NDUFV2	24 kDa, Nqo2, NuoE	E	217	5-216	97.7	2Fe2S
NDUFS1	75 kDa, Nqo3, NuoG	G	704	6-693	97.7	2Fe2S, 2 x 4Fe4S
NDUFS2	49 kDa, Nqo4, NuoCD	D	430	1-430	100	Dimethyl-Arg85. No sidechains 40-42 Deactive: 52-60 absent
NDUFS3	30 kDa, Nqo5, NuoCD	C	228	8-213	90.4	
NDUFS7	PSST, Nqo6, NuoB	B	189	35-189	82.0	4Fe4S
NDUFS8	TYKY, Nqo9, NuoI	I	178	1-178	100	No side-chains 8-14, 2 x 4Fe4S
ND1	Nqo8, NuoH	H	318	1-318	100	N-formyl. Deactive: 204-215 absent
ND2	Nqo14, NuoN	N	345	1-344	99.7	N-formyl
ND3	Nqo7, NuoA	A	115	1-115	100	N-formyl. Deactive: 28-48 absent
ND4	Nqo13, NuoM	M	459	1-459	100	N-formyl
ND4L	Nqo11, NuoK	K	98	1-98	100	N-formyl
ND5	Nqo12, NuoL	L	607	1-606	99.8	N-formyl
ND6	Nqo10, NuoJ	J	172	1-171	99.4	N-formyl Deactive: no sidechains 84-118
NDUFV3	10 kDa	s	69	29-68	58.0	
NDUFS4	18 kDa	Q	133	9-133	94.0	
NDUFS5	15 kDa	e	105	1-104	99.0	2 x Cys-Cys
NDUFS6	13 kDa	R	96	1-95	99.0	Zn^{2+}

NDUFA1	MWFE	a	70	1-67	95.7	
NDUFA2	B8	S	98	13-95	84.7	
NDUFA3	B9	b	83	1-83	100	N-acetyl
NDUFA5	B13	V	115	4-115	97.4	
NDUFA6	B14	W	130	17-130	87.7	
NDUFA7	B14.5a	r	112	1-76, 90-122	89.3	N-acetyl
NDUFA8	PGIV	X	171	1-171	100	4 x Cys-Cys
NDUFA9	39 kDa	P	342	1-342	100	Deactive: 186-196, 251-278, 324-335 absent, NADPH
NDUFA10	42 kDa	O	320	1-320	100	ADP
NDUFA11	B14.7	Y	140	1-140	100	Cys-Cys
NDUFA12	B17.2	q	145	1-144	99.3	N-acetyl
NDUFA13	B16.6	Z	143	5-143	97.2	
NDUFAB1α	SDAPα	T	88	8-82	85.2	4'-phosphopantethine + 3-hydroxyundecanoate
NDUFAB1β	SDAPβ	U	88	1-88	100	4'-phosphopantethine + 3-hydroxydodecanoate
NDUFB1	MNLL	f	56	4-54	91.1	
NDUFB2	AGGG	j	72	6-65	83.3	
NDUFB3	B12	k	103	20-91	69.9	
NDUFB4	B15	m	128	5-128	96.9	
NDUFB5	SGDH	h	143	6-142	95.8	
NDUFB6	B17	i	127	1-35, 62-127	79.5	N-acetyl. No sidechains 62-65, 124-126. 36-61 absent
NDUFB7	B18	o	136	12-116	77.2	2 x Cys-Cys
NDUFB8	ASHI	l	157	3-154	96.8	
NDUFB9	B22	n	178	1-176	98.9	
NDUFB10	PDSW	p	175	2-171	97.1	2 x Cys-Cys
NDUFB11	ESSS	g	122	21-121	82.7	
NDUFC1	KFYI	c	49	2-48	95.9	
NDUFC2	B14.5b	d	120	1-119	99.2	

Table 4.1: Summary of the models for the subunits of mouse complex I.

	Active	Deactive
Data Collection		
Pixel size (Å)	1.05	1.33
Defocus range (μm)	2.2-3.4	2.2-3.4
Voltage (kV)	300	300
Number of particles in refinement	20,370	17,066
Estimated accuracy of rotations (º)	0.47	0.53
Map refinement		
Resolution (Å)	3.30	3.93
B-factor used for map sharpening (Å^2)	-63	-62
Model statistics		
Non-hydrogen atoms	66,812	64,925
Protein residues	8,164	8,072
% of total	95.9	94.8
Core subunit residues	4,408	4,366
% of total	97.9	96.9
Supernumerary subunit residues	3,756	3,718
% of total	93.6	92.3
Residue with sidechains	8,148	8,020
% of total	95.7	94.2
RMS deviations		
Bonds (Å)	0.0071	0.0130
Angles (°)	1.36	1.36
Validation		
Molprobity score	2.08	2.19
Clashscore, all atoms	7.76	8.95
EMringer score	3.56	2.10
Ramachandran plot		
Favored (%)	85.3	81.8
Outliers (%)	0.22	0.45

Table 4.2: Data collection, refinement, and model statistics for the structures of active and de-active mouse complex I.

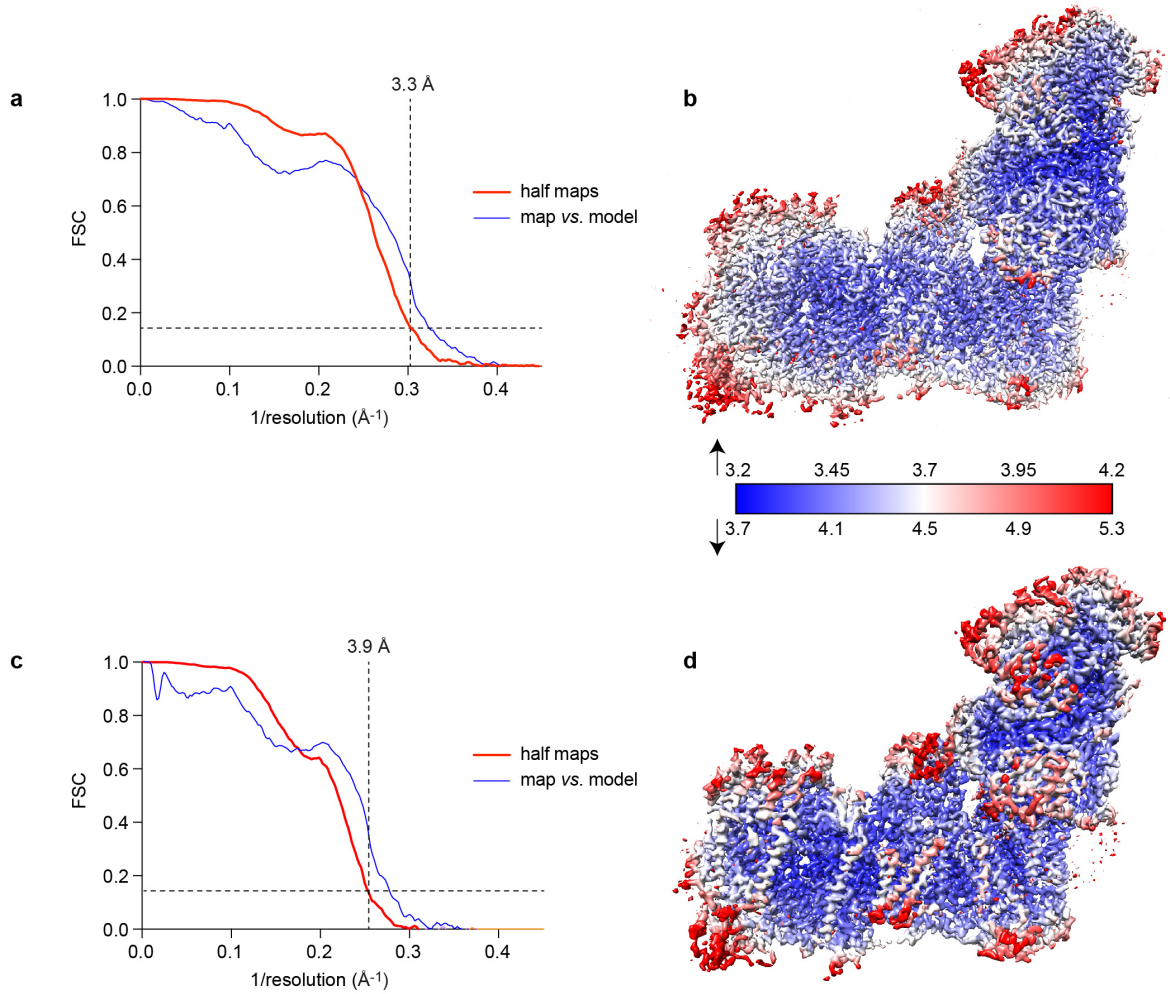


Figure 4.11: Resolution estimations of the maps for the active and de-active states of mouse complex I. a, b: data on the active complex. c, d: data on the deactive complex. The estimated resolutions, defined where the line at FSC = 0.143 crosses the red curve, are 3.3 Å for the active complex (a) and 3.9 Å for the deactive complex (c). In both cases the refined models agree well with the maps, as shown by the map vs. model FSC curves (blue). In b) and d), local resolutions were estimated using the Local Resolution function in RELION with default parameters and plotted using UCSF Chimera.

4.5 Local resolution of the cryo-EM maps of mouse complex I

The local resolution of both the active and de-active complex I cryo-EM maps varies throughout the structure, particularly between the core and supernumerary subunits (Figure 4.13). The central core subunits are well-resolved, as evident by their relatively high resolution,

whilst the tip of the membrane arm (ND5) and matrix arm (NDUFV1, NDUFV2 and the large domain of NDUFS1) have relatively poor resolution and therefore are not well-resolved (Figure 4.12). The resolution decrease towards the end of each core domain can be explained well by considering the surface area contact each subunit makes with the rest of the protein (Table 4.3). Noticeably, the relative surface area contact of each subunit generally coincides with its local resolution, meaning proteins that have relatively few surface contacts with the rest of the complex, such as ND5 (49%), NDUFV2 (36%), NDUFV1 (31%) and the large domain of NDUFS1 (30%) display a lower resolution compared to the rest of the core subunits, which on average has relatively more surface area contact 69 ± 6% (Table 4.3). This may be because a lower surface contact area of a given protein with the rest of the complex results in a distribution of orientations for that protein compared to the rest of the structure, thus displaying a lower resolution because of its flexibility. Similarly, the supernumerary subunits display a lower resolution than the core subunits, again owing to the relatively low contact areas of the supernumerary subunits (42 ± 12%) compared to the rest of the complex. Interestingly, the resolution is asymmetric across the membrane domain of complex I instead of the extremities of the complex, which is common in image processing. This suggests that the surface-area effect is substantial and variations are not due to imperfect particle alignment (Figure 4.12). The origin of the resolution asymmetry is unclear, although the better-resolved side contains the ND5 transverse helix and contacts with complex III in the respirasome (Letts et al., 2016a; Wu et al., 2016a), hinting that inter-complex contacts may have affected rigidity. Finally, supernumerary subunits distant from the core tend to have particularly low local resolutions. For example, subunits NDUFA2 and NDUFAB1α, which contact only the large domain of NDUFS1 (25% contact area) and NDUFA6 (15% contact area), respectively are clearly noticeable in Figure 4.12, and NDUFC1, which is bound peripherally to the complex by supernumerary subunit NDUFC2, has the lowest local resolution estimate of any subunit (Table 4.3). The concept of low surface area contacts producing decreased resolution through increased structural dispersion may thus help to explain how the resolution varies through the complex, particularly for the supernumerary subunits, which rely less on tight and specific interactions with the complex for their structure/function than the core subunits.

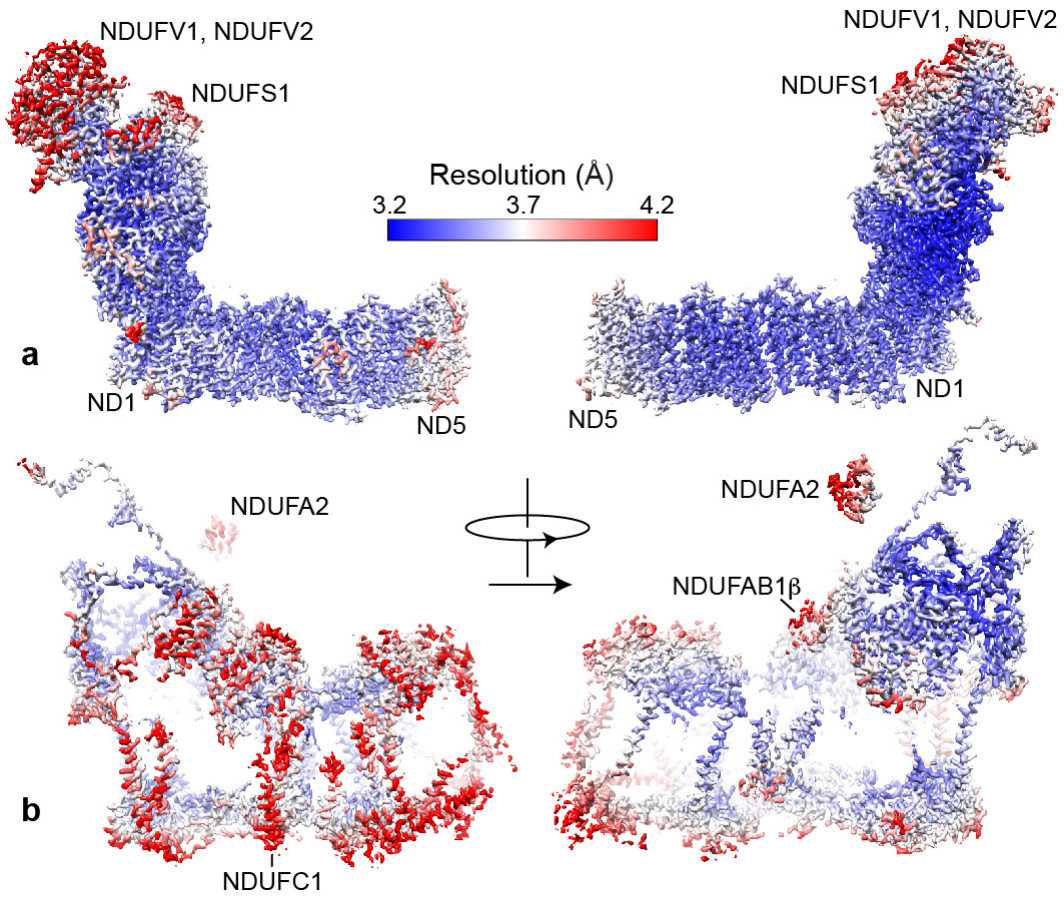


Figure 4.12: The resolution of densities for the core and supernumerary subunits in the 3.3 Å resolution map of active mouse complex I. Local resolutions were estimated using the Local Resolution function in RELION with default parameters. The figure was created using UCSF Chimera, using the model to separate the core (a) and supernumerary (b) subunit densities.

Subunit	Surface area			Local Resolution (Å)	Subunit contacts in order of decreasing contact area	
	Total (Å^2)	Buried (Å^2)	% Buried		Core subunits	Supernumerary subunits
NDUFV1	18,800	6,100	33	3.9 ± 0.4	NDUFV2, NDUFS1	NDUFV3, NDUFS4, NDUFA7
NDUFV2	13,600	4,900	36	3.7 ± 0.3	NDUFV1, NDUFS1	NDUFV3, NDUFS4, NDUFS6
NDUFS1	28,500	10,100	35	3.6 ± 0.2	NDUFV1, NDUFS2, NDUFV2, NDUFS3, NDUFS8	NDUFS4, NDUFA2, NDUFA12, NDUFA6, NDUFA7, NDUFS6, NDUFA9, NDUFV3
NDUFS1 (small)	12,200	7,100	58	3.6 ± 0.2	(NDUFS1 large), NDUFV1, NDUFS2, NDUFV2, NDUFS3, NDUFS8	NDUFS4, NDUFA7, NDUFS6, NDUFA12, NDUFV3
NDUFS1 (large)	19,100	5,600	30	3.5 ± 0.3	NDUFS1 (small), NDUFS3, NDUFS8	NDUFA2, NDUFS4, NDUFA12, NDUFA6, NDUFA9, NDUFV3
NDUFS2	21,300	14,400	68	3.4 ± 0.2	NDUFS3, NDUFS8, NDUFS7, ND1, ND2, NDUFS1, ND3, ND4, ND4L, ND5	NDUFA7, NDUFA13, NDUFA10, NDUFA5, NDUFS6, NDUFB11, NDUFA6, NDUFS4, NDUFA12

NDUFS3	14,500	10,200	70	3.4 ± 0.1	NDUFS2, NDUFS7, NDUFS1, NDUFS8, ND3	NDUFA5, NDUFS4, NDUFA7, NDUFA6, NDUFA9
NDUFS7	9,000	6,300	70	3.3 ± 0.1	NDUFS2, ND1, NDUFS8, NDUFS3, ND3	NDUFA9, NDUFA12, NDUFA6, NDUFA7, NDUFS4
NDUFS8	13,100	9,600	73	3.4 ± 0.0	NDUFS2, NDUFS7, ND1, NDUFS1, NDUFS3	NDUFA12, NDUFS6, NDUFA7, NDUFA13, NDUFA3, NDUFS4, NDUFA1, NDUFA9, NDUFA5
ND1	16,000	10,700	67	3.5 ± 0.1	ND3, NDUFS7, NDUFS2, ND6, NDUFS8	NDUFA1, NDUFA13, NDUFA3, NDUFA8, NDUFA9, NDUFA12, NDUFA7
ND2	15,200	9,800	65	3.4 ± 0.1	ND4L, ND4, ND5, NDUFS2, ND6, ND3	NDUFA10, NDUFC2, NDUFS5, NDUFA11, NDUFB5, NDUFA8, NDUFB11
ND3	11,000	7,400	67	3.5 ± 0.1	ND1, ND6, NDUFS2, NDUFS7, ND4L, ND2, NDUFS3	NDUFA3, NDUFA6, NDUFA9, NDUFA13, NDUFS5, NDUFA1
ND4	18,200	11,000	60	3.5 ± 0.1	ND5, ND2, NDUFS2	NDUFB11, NDUFB4, NDUFB5, NDUFB1, NDUFB10, NDUFB8, NDUFB9, NDUFA8, NDUFC2, NDUFA11, NDUFA10
ND4L	7,000	5,700	83	3.4 ± 0.1	ND6, ND2, ND5, ND3, NDUFS2	NDUFS5
ND5	28,600	13,900	49	3.6 ± 0.1	ND4, ND2, ND4L, NDUFS2	NDUFB8, NDUFB2, NDUFB9, NDUFB10, NDUFB6, NDUFB4, NDUFB7, NDUFB3, NDUFA11, NDUFAB1β, NDUFB5, NDUFB11
ND6	13,000	8,300	64	3.5 ± 0.1	ND4L, ND3, ND1, ND2	NDUFA13, NDUFS5, NDUFA1, NDUFA9, NDUFA3, NDUFA6
NDUFS4	10,800	6,500	60	3.4 ± 0.1	NDUFS1, NDUFS3, NDUFV1, NDUFS8, NDUFV2, NDUFS2, NDUFS7	NDUFA6, NDUFA9, NDUFV3, NDUFA12, NDUFS6, NDUFA5
NDUFS5	9,600	5,400	56	3.6 ± 0.2	ND2, ND6, ND4L, ND3	NDUFA13, NDUFB5, NDUFA8, NDUFC2, NDUFA1, NDUFA3
NDUFS6	7,500	3,000	40	3.4 ± 0.1	NDUFS8, NDUFS1, NDUFS2, NDUFV2	NDUFA9, NDUFA12, NDUFS4
NDUFA1	6,200	3,700	60	3.6 ± 0.1	ND1, ND6, NDUFS8, ND3	NDUFA8, NDUFA13, NDUFS5, NDUFA7, NDUFA12
NDUFA2	5,500	1,400	25	4.2 ± 0.4	NDUFS1	-
NDUFA3	8,100	3,300	41	4.1 ± 0.4	ND1, ND3, NDUFS8	NDUFA8, NDUFA13, NDUFS5, ND6, NDUFB5
NDUFA5	8,500	2,900	34	3.9 ± 0.3	NDUFS3, NDUFS2, NDUFS8	NDUFA7, NDUFA10, NDUFS4, NDUFA6
NDUFA6	9,300	4,500	48	3.5 ± 0.1	NDUFS3, NDUFS1, ND3, NDUFS2, NDUFS7, ND6	NDUFAB1α, NDUFS4, NDUFA9, NDUFA5
NDUFA7	11,500	6,000	52	3.8 ± 0.3	NDUFS2, NDUFS3, NDUFS8, NDUFS1, NDUFV1, ND1, NDUFS7	NDUFA12, NDUFA13, NDUFA5, NDUFA1
NDUFA8	14,100	6,500	46	3.8 ± 0.2	ND2, ND4, ND1	NDUFA13, NDUFA3, NDUFA1, NDUFC2, NDUFB5, NDUFS5, NDUFC1, NDUFB1
NDUFA9	17,000	4,100	24	3.5 ± 0.2	NDUFS7, NDUFS3, ND1, ND3, NDUFS1, ND6, NDUFS8	NDUFS6, NDUFA6, NDUFS4, NDUFA12
NDUFA10	17,000	3,900	23	3.8 ± 0.3	ND2, NDUFS2, ND4	NDUFC1, NDUFB11, NDUFC2, NDUFA5
NDUFA11	9,300	1,600	17	3.6 ± 0.1	ND5, ND2, ND4, NDUFS2	NDUFB5, NDUFB4
NDUFA12	11,800	4,600	37	3.5 ± 0.2	NDUFS8, NDUFS1, NDUFS7, ND1, NDUFS2	NDUFS6, NDUFA7, NDUFS4, NDUFA9, NDUFA1
NDUFA13	14,000	8,200	59	3.7 ± 0.2	ND3, NDUFS2, ND1, NDUFS8, ND6	NDUFA1, NDUFA3, NDUFS5, NDUFB5, NDUFA7, NDUFA8
NDUFAB1α	5,400	900	17	4.0 ± 0.2	-	NDUFA6
NDUFAB1β	6,300	3,300	53	3.9 ± 0.2	ND5	NDUFB9, NDUFB3, NDUFB6, NDUFB2, NDUFB8, NDUFB5
NDUFB1	5,600	1,700	31	4.1 ± 0.4	ND4	NDUFB5, NDUFB10, NDUFA8, NDUFC2
NDUFB2	6,700	3,200	48	4.1 ± 0.2	ND5	NDUFB3, NDUFB7, NDUFAB1β, NDUFB8, NDUFB6
NDUFB3	7,000	2,800	39	4.1 ± 0.3	ND5	NDUFB2, NDUFB9, NDUFAB1β
NDUFB4	12,600	5,900	47	3.6 ± 0.2	ND4, ND5	NDUFB8, NDUFB9, NDUFB10, NDUFB11, NDUFC2, NDUFA11

NDUFB5	13,700	7,700	56	3.9 ± 0.3	ND4, ND2, ND5	NDUFB10, NDUFS5, NDUFB11, NDUFC2, NDUFB1, NDUFA8, NDUFB9, NDUFB6, NDUFA11, NDUFA13, NDUFAB1β, NDUFA3
NDUFB6	10,700	5,000	47	4.1 ± 0.3	ND5	NDUFB10, NDUFB9, NDUFB7, NDUFAB1β, NDUFB5, NDUFB2, NDUFB11
NDUFB7	9,500	3,900	41	4.1 ± 0.3	ND5	NDUFB8, NDUFB6, NDUFB2, NDUFB10
NDUFB8	13,200	5,500	42	3.6 ± 0.2	ND5, ND4	NDUFB4, NDUFB7, NDUFB9, NDUFB10, NDUFAB1β, NDUFB2
NDUFB9	15,300	7,100	47	3.8 ± 0.3	ND5, ND4	NDUFAB1β, NDUFB4, NDUFB6, NDUFB3, NDUFB5, NDUFB8, NDUFB11
NDUFB10	15,400	8,000	52	4.0 ± 0.4	ND5, ND4	NDUFB11, NDUFB6, NDUFB5, NDUFC2, NDUFB7, NDUFB4, NDUFB1, NDUFB8
NDUFB11	10,500	5,300	51	3.8 ± 0.3	ND4, NDUFS2, ND5, ND2,	NDUFB10, NDUFB5, NDUFA10, NDUFC2, NDUFB9, NDUFB4, NDUFB6
NDUFC1	5,100	1,500	29	4.3 ± 0.3	-	NDUFC2, NDUFA10, NDUFA8
NDUFC2	10,500	4,900	47	3.8 ± 0.2	ND2, ND4	NDUFC1, NDUFB5, NDUFB10, NDUFA8, NDUFA10, NDUFS5, NDUFB11, NDUFB4, NDUFB1
NDUFV3	4,800	2,600	53	3.7 ± 0.2	NDUFV1, NDUFV2, NDUFS1	NDUFS4

Table 4.3: The interfaces between subunits in the 3.3 Å structure of mouse complex I. Surface areas of contacts between subunits were calculated using the PDBePISA server (http://www.ebi.ac.uk/pdbe/pisa/) and reported to 100 Å^2 accuracy. Local resolutions were estimated using the Local Resolution function in RELION with default parameters, then UCSF Chimera was used to extract the resolution at each atomic coordinate in each subunit and the values were averaged.

4.6 Assignment of the 3.3 Å resolution structure to the active state, and comparison with a de-active structure

The 3.3 Å resolution structure of mouse complex I, determined from the protein as isolated, was assigned to the active state for the following reasons. First, the global RMSD values calculated in Pymol, which gives an indication of similarity between two atomic coordinate models, were much lower when the active mouse complex I model was compared with the active (class 2) bovine complex I model than the de-active (class 1) or broken (class 3) models (1.6, 3.2 and 4.2 Å, respectively). Second, the different orientations of the hydrophilic and membrane domains in the active and de-active states are captured visually by the different positions of NDUFA5 (hydrophilic domain) and NDUFA10 (membrane domain), showing clearly that the mouse structure matches the active bovine structure (Figure 4.13a). Third, densities for all the loops that become disordered in the de-active state are clearly visible in the

mouse map (Blaza et al., 2018; Zhu et al., 2016), and remained so even when its resolution was decreased to 4.3 Å, to match the resolution of the original bovine structures.

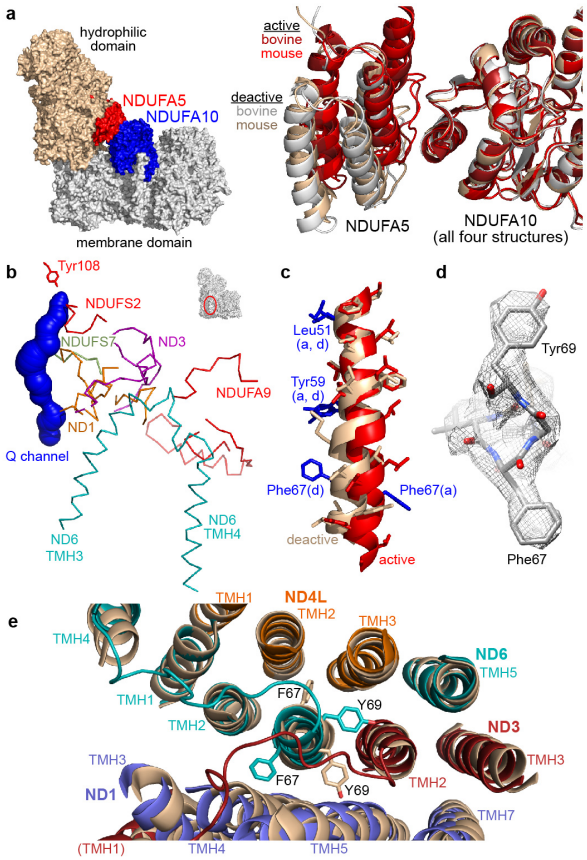


Figure 4.13: Differences in structure between the active and de-active states of mouse complex I. (a) Left - NDUFA5, on the hydrophilic domain, and NDUFA10, on the membrane domain, form a second contact between the two domains. Right – the different arrangements of NDUFA5 and NDUFA10 are illustrated by superimposing the NDUFA10 subunit in the NDUFA5-NDUFA10 pairs from the bovine and mouse de-active and active states. NDUFA5 is present in markedly different positions in the active (red) and de-active (wheat/grey) states. (b) Structural elements that become disordered or change between the active and de-active states. The ubiquinone binding channel determined by CAVER 3.0 is shown for reference (Chovancova et al., 2012). (c) The different structures of ND6-TMH3 in the active (red) and de-active (wheat) states, superimposed on the N-terminal section, and the positions of key sidechains. The π-bulge is present only in the de-active state. (d) Densities for key sidechains in the active state. (e) Rotation of the bulky sidechains of Phe67 and Tyr69 between the TMHs of ND3, ND4L and ND1. The structures were superimposed on ND4L with the active structure in different colors and the de-active in wheat. Only the TMHs are shown, plus the loops between ND3-TMHs 1 and 2 and ND6-TMHs 3 and 4.

To confirm that the active and de-active structural states observed for bovine complex I apply equally to the mouse enzyme, mouse complex I was prepared in the deactivate state using the method of Blaza *et al.,* (Blaza et al., 2018), and its structure was determined to 3.9 Å resolution (Figure 4.11), following 3D particle classification into a major (69%) and a minor (25%) class (Figure 4.8b). Comparisons of the minor class with the three classes of bovine complex I suggested it represents the active state of the mouse enzyme (Zhu et al., 2016). The complex I model for the major class exhibited its lowest global RMSD value when compared with the de-active bovine complex I, with values of 1.6, 2.4 and 2.7 Å for classes 1, 2 and 3, respectively. Moreover, Figure 4.13a reveals a close match in the relative positions of NDUFA5 and NDUFA10 between the de-active bovine and mouse enzyme. Consistent with de-active bovine complex I, there are no clear densities are observed for the loops of ND1 TMH5-6, ND3 TMH1-2 and NDUFS2 β1-β2, or several loops in NDUFA9 (Table 4.1). However, key features which are absent in the inactive state (class 3) such as the transverse helix of ND5 are clearly visible, which means the deactivation of the enzyme has not denatured it (Zhu et al., 2016). Therefore, the structural changes accompanying deactivation are common to both the bovine and mouse enzyme, providing a clear understanding of the active-deactive transition seen in mammalian complex I.

Figure 4.13a shows that NDUFA5 and NDUFA10 form an interface between the membrane arm and the hydrophilic arm of complex I, which is the only interface between the two domains other than the main domain contact between the core subunits. There are changes in the interface interaction between NDUFA5/NDUFA10 upon deactivation. In the active state, the buried interface between the two subunits is 400 $Å^2$, including three salt-bridges and three hydrogen bonds. In contrast, in the de-active state, the interface is decreased to 200 $Å^2$ and includes only one salt-bridge and one hydrogen bond. The implication that NDUFA5/NDUFA10 interactions stabilise the active state is supported by coarse-grained simulations that illustrated how NDUFA5 and NDUFA10 restrict global vibrational modes in the mammalian complex that resemble the changes required to interconvert the active and de-active states (Di Luca and Kaila, 2018).

In addition to the loops in ND1, ND3, NDUFS2 and NDUFA9, which are known to become disordered upon deactivation (Blaza et al., 2018; Zhu et al., 2016a), the higher-resolution active and de-active mouse structures reveal subtle differences in ND1, ND6 and NDUFS7,

illustrating better how the disorder propagates from the ubiquinone-binding site (Figure 4.13b). In ND6, the TMH3-4 loop is affected in the active-de-active transition (i.e. becomes more disordered), and in the de-active map, TMH4 is poorly resolved. Furthermore, a striking difference between the active and de-active states is revealed in ND6-TMH3. In all existing complex I structures, a π-bulge (a stretch of a π-helix) has been modelled in ND6-TMH3, following its identification in the high-resolution bacterial membrane domain of complex I (Efremov and Sazanov, 2011). The π-bulge is present in de-active mouse complex I, but not in the active mouse enzyme. Consequently, in the active state TMH3 is extended, and the residues in its C-terminal section are rotated around the helical axis (Figure 4.13c). Although ND6-Tyr59, which may form a functionally-important interaction with ND4L-Glu34, is at the start of the π-bulge and so not affected, two aromatic residues, Phe67 and Tyr69, undergo large changes (Figure 4.13d). They rotate, respectively, from between ND4L-TMH2/3 and ND1-TMH5/ND3-TMH2 in the de-active state to between ND1-TMH4/5 and ND3-TMH2/ND4L-TMH3 in the active state (Figure 4.13e). The rotation locks the enzyme into different conformations, explaining why ND6 moves relative to ND1 and ND3 between the active and de-active states (Zhu et al., 2016), and thus confirming the importance of the ND6 π-bulge for catalysis.

4.7 Phospholipid densities

Densities for twenty phospholipids were identified in the map for active mouse complex I (Figure 4.14 and Table 4.4). Four of them are cardiolipins, 11 phosphatidylethanolamines and five phosphatidylcholines. Many of the headgroup phosphates are stabilized by arginines, histidines or tyrosines (Hunte, 2005). Eight phospholipids overlap with the 12 phospholipids that were previously modelled in ovine complex I (Fiedorczuk et al., 2016), and nine and 11 of phospholipids modelled in the mouse structure overlap with the 11 and 18 phospholipids modelled in the porcine and human respirasomes (Guo et al., 2017; Wu et al., 2016), respectively (Table 4.4). Altogether nine cardiolipins have been modelled, corresponding to the ~10 cardiolipins that co-purify with bovine complex I, along with a similar number of interchangeable phosphatidylethanolamines and phosphatidylcholines.

The locations of the phospholipids (Figure 4.14 and Table 4.4) suggest they stabilize intra-protein and subunit interactions. Four stabilize the C-terminus of ND5, one at its interface with

NDUFA11 and ND2 and two 'behind' its transverse helix, and on the heel, they stabilize hydrophilic-domain amphipathic helices and supernumerary TMHs against ND1 (Figure 4.14b). Many phospholipids are packed between the core membrane subunits. Notably, the contact areas between ND2, ND4 and ND5 are low, with less than 20% of the surface of ND4 buried at interfaces with ND2 and ND5, suggesting loose contacts between the core subunits of the membrane domain. Therefore, along with the 'cage' of supernumerary subunits, phospholipids may also be crucial for the integrity of the membrane domain. Finally, two phospholipids (Figure 4.14c) interact with NDUFA9 residues relevant to the active to de-active transition. In particular, one of them extends from close to the NAD(P)H ribose-phosphate, along an interface between NDUFS7 and two regions (a loop in NDUFA9 and the start of the ND3-TMH1-2 loop) that become disordered in the de-active state. Therefore, protein-phospholipid interactions may stabilize loops that are inherently unstable and susceptible to becoming disordered.

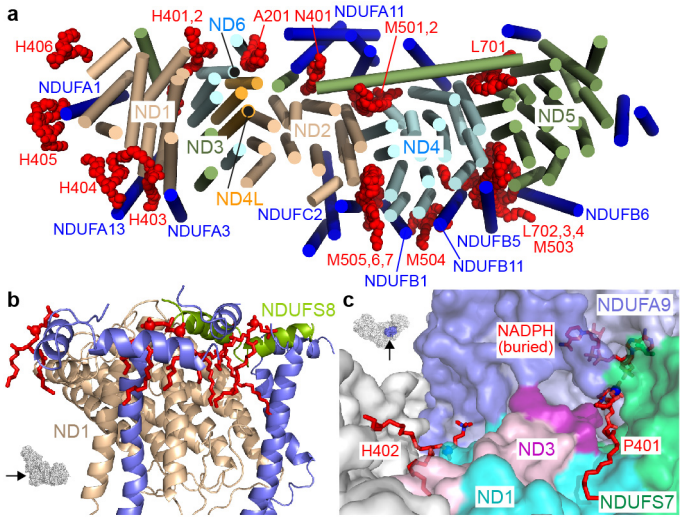


Figure 4.14: Phospholipids in the 3.3 Å resolution structure of mouse complex I. a) Membrane-bound helices in the complex I viewed from the matrix side, with phospholipids in red and labelled with the nomenclature in Supplementary Table 4. b) Phospholipids stabilizing amphipathic helices on the matrix side of the heel of the complex. The structure shown in blue is from supernumerary subunits NDUFA1, NDUFA3, NDUFA7, NDUFA12 and NDUFA13 and phosphorus atoms are marked with spheres. c) Two phospholipids that interact with NDUFA9 may stabilize loops in ND3 and NDUFA9 that become disordered in the de-active state (shown in darker colours than the rest of the two subunits). The arrows on the insets to b) and c) indicate the directions of view.

Phospholipid in Mouse	Confidence	Side	Location	Human (5XTD)	Ovine (5LNK)	Porcine (5GUP)
H-401 3PE	1	IMS	Interface of ND1, ND3-TMH1 and ND6 (also loops of NDUFA1 and NDUFA13 in IMS).	PE	PE	PE
H-402 PC1	2	Matrix	Interface of ND1 and ND6 (also loops in C-terminal domain of NDUFA9).	-	-	-
A-201 3PE	2	Matrix	Interface of ND4L and ND6 (also C-terminal TMH of ND5).	-	-	-
N-401 3PE	2	Matrix	Interface of ND2, NDUFA11 and transverse and final TMH of ND5.	CDL	-	CDL
M-501 3PE	1	Matrix	Interface of ND2, ND4 and ND5 transverse helix (also N-terminus of NDUFS2). See also M-502.	PE	-	PE
M-502 3PE	2	IMS	Interface of ND2, ND4 and NDUFA11. See also M501.	PE	-	PE
L-701 3PE	1	Matrix	Interface of ND4 and ND5 and ND5 transverse helix (also loops and TMH of NDUFB8).	PE	-	PE
L-702 CDL	1	Matrix	Interface of ND4, ND5 and NDUFB5 (also NDUFB11). See also L-703, M-503, L-704.	CDL	CDL	CDL
L-703 CDL	2	IMS	Interface of ND4, ND5, NDUFB5, NDUFB6 and NDUFB11 (also NDUFB10 on IMS face). See also L-702, M-503, L-704.	CDL	CDL	-
M-503 3PE	1	IMS	Interface of ND4, ND5 and NDUFB11. See also L-702, L-703, L-704.	PC	(CDL)	-
L-704 3PE	1	IMS	Interface of ND5, NDUFB5 and NDUFB6. See also L-702, L-703, M-503.	-		-
M-504 3PE	2	IMS	Interface of ND4, NDUFB1 and NDUFB11 (also NDUFB5 on IMS face).	-	-	-
M-505 3PE	2	Matrix	Interface of ND2, ND4 and NDUFC2 (also C-terminal helix of NDUFA10). See also M-506, M-507.	CDL	-	CDL
M-506 3PE	2	Matrix	Interface of ND2, ND4 and NDUFC2 (also NDUFB1 and C-terminal loop of NDUFA10). See also M-505, M-507.		(PE)	
M-507 CDL	1	IMS	Interface of ND2, ND4, NDUFB1 and NDUFC2 (also C-terminus of NDUFA8 on IMS face). See also M-505, M-506.	CDL	PC	CDL
H-403 PC1	1	Matrix	Interface of ND1, NDUFA3 and NDUFA13 (also helix in NDUFS2 and N-terminus of NDUFS8). See also H-404.	PC	PE	PC
H-404 PC1	1	Matrix	Interface of ND1 and NDUFA13 (also helices of NDUFS8 N-terminus and NDUFA7). See also H-403.	-	-	-
H-405 CDL	1	Matrix	Interface of NDUFA1 and helix of NDUFA12 (also ND1, NDUFS8 and NDUFA7).	-	CDL	-
H-406 PC1	2	Matrix	Interface of helices on ND1, NDUFS7 and NDUFA12 (also NDUFS8).	-	-	-
P-401 PC1	2	Matrix	Interface of ND3-TMH1, amphipathic helices on NDUFS7 and loops in ND3, NDUFS7 and NDUFA9.	-	-	-

Table 4.4: Phospholipids modelled in the structure of active mouse complex I. The phospholipids were assigned on two confidence levels, where 1 is the highest, based on the quality of the density and interactions with nearby residues. Cases in which partial overlap between phospholipids are observed in different structures are indicated with brackets. Interactions of phospholipids with subunits were determined with a 5 Å cut-off.

4.8 Nucleotide binding to NDUFA10

The improved resolution of the active mouse structure has allowed the tetrahedral coordination of the Zn in NDUFS6 and the acyl chains in NDUFAB1α and NDUFAB1β to be clearly visualized, revealed the coordination of the NAD(P)H in NDUFA9 by conserved charged residues, confirmed 11 disulfide bonds on the inter-membrane space side of the complex and showed that no disulfide is present in the thioredoxin-like subunit NDUFA2 (located in the reducing matrix environment). Most strikingly, however, it has revealed clear density for a bound nucleotide in subunit NDUFA10, a member of the nucleoside-kinase family (Figures 4.15a and b) that transfer a phosphate from a nucleotide triphosphate to a nucleoside (Eriksson et al., 2002). The density has been tentatively modelled as ADP because it resembles a purine nucleoside plus two phosphates, and ADP is abundant in the matrix. The putative ADP is coordinated by the sidechains of widely conserved residues: the adenosine by Tyr83, Gln97, Arg104 (which interacts with Glu53 and Arg126), Phe134, Glu191 and Tyr198, and the phosphates by Lys35 (P-loop), Glu125 and Lys186. However, comparison with structural information from canonical nucleoside kinases, such as ADP-bound human deoxycytidine kinase (Sabini et al., 2008), showed that the ADP is bound with its adenosine in the nucleoside-binding site and its two phosphates towards the P-loop nucleotide-phosphates binding site (Figure 4.15d). Therefore, it is bound in a configuration resembling the feedback-inhibition mode characterized for deoxythymidine triphosphate bound to *Drosophila melanogaster* deoxyribonucleoside kinase (Mikkelsen et al., 2003).

Similar, less well-defined densities are present (but were not interpreted) in NDUFA10 in the maps for the human and porcine respirasomes (Guo et al., 2017; Wu et al., 2016a), but not in the maps for ovine or bovine complex I (Fiedorczuk et al., 2016; Zhu et al., 2016a). In the bovine maps increased density in the ADP phosphate-binding region can be attributed to phosphorylation of Ser36 (Schilling et al., 2005) (Figure 4.15c), which has now been confirmed in the structurally-characterized (bovine) preparation (Figure 4.16). Ser36 is conserved in ovine complex I, but substituted by Asn/Gly in the mouse, porcine and human homologs, suggesting phosphorylation blocks nucleotide binding. As the active site is accessible to the solution and key residues are present (Elurbe and Huynen, 2016), phosphorylation of Ser36 and inhibitory substrate binding may together explain why no nucleoside kinase activity has yet been reported for mammalian complex I.

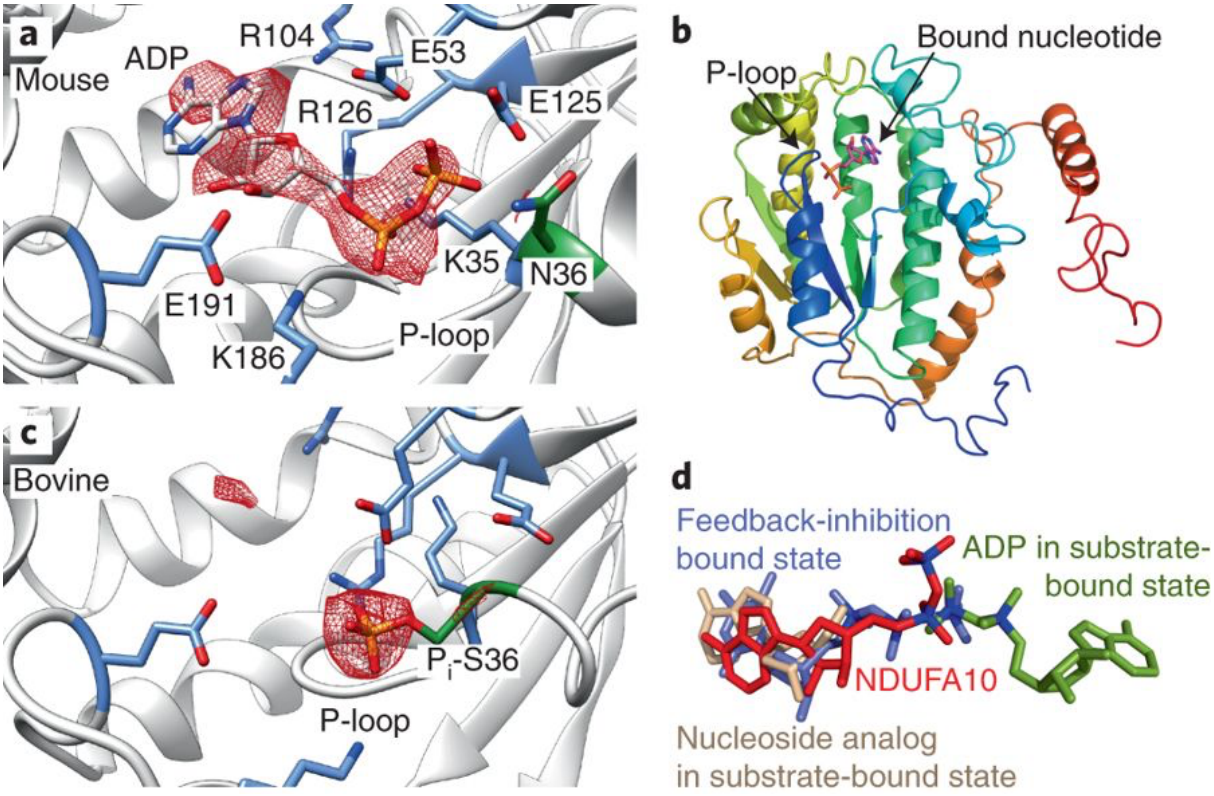


Figure 4.15: Nucleotide/nucleoside binding to NDUFA10. (a) Density for the bound nucleotide/nucleoside (modelled as ADP) in NDUFA10 in active mouse complex I. The density has been carved around the modelled ADP and the phosphate in the overlaid model from panel c (2 Å carve radius, threshold level of 0.075). Key residues that interact with the nucleotide/nucleoside are indicated, and N36, the homolog of S36 is in green. (b) Overview of the typical nucleoside-kinase fold of NDUFA10 with nucleotide bound and the P-loop adjacent to the bound phosphates. (c) The nucleotide/nucleoside-binding site in the bovine complex (taken from 5O31.PDB (Blaza et al., 2018)) with a phosphate added to the side-chain of Ser36 (green) on the basis of mass spectrometry evidence. The density has been carved around the phosphate and the ADP in the overlaid model from panel a (2 Å carve radius, threshold level of 0.16). (d) Comparison of the binding mode of the ADP modelled in NDUFA10 (red) with deoxythymidine triphosphate bound in the feedback inhibition mode in *Drosophila melanogaster* deoxyribonucleoside kinase (1OE0.PDB (Mikkelsen et al., 2003), blue) and with ADP (green) and a nucleoside analogue (wheat) in the substrate-binding site of human deoxycytidine kinase (2ZI4.PDB (Sabini et al., 2008)). The structures were overlaid using the whole nucleoside kinase fold structures.

a

	Measured mass	Calculated mass	Difference
NDUFA10-Pi	36,790.7 Da	36,787.0 Da	3.7 Da
NDUFA10	36,707.3 Da	36,707.0 Da	0.3 Da

b

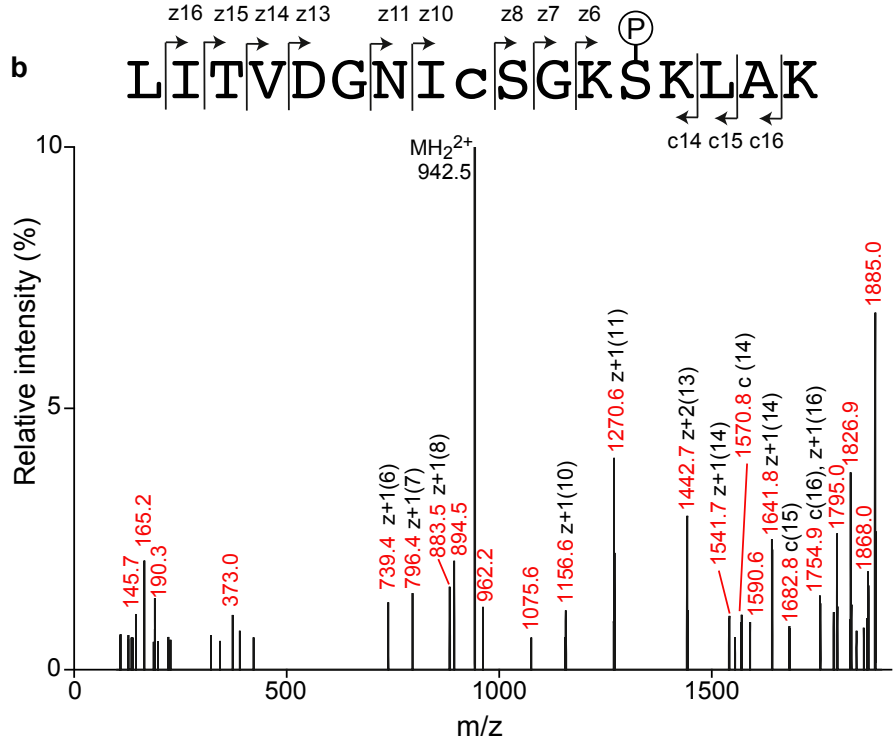


Figure 4.16: Mass spectrometry evidence for the phosphorylation of bovine NDUFA10-Ser36. (a) Intact mass measurements revealed two masses for NDUFA10, for the unmodified and singly phosphorylated versions (present at the point of measurement). (b) Spectrum of fragment ions produced by ETD of a doubly charged precursor ion 942.48 (m/z) with the precursor ion truncated to 10% relative intensity. The observed neutral peptide mass for the phosphorylated peptide is 1882.955 Da, relative to a calculated mass of 1882.959 Da. The observed peptides are mapped onto the amino acid sequence where the phosphorylation site is marked and c is carbamidomethylcysteine. The z7 and z8 fragments confirm Ser36, rather than Ser33, as the site of phosphorylation.

4.9 The second coordination sphere of cluster N2

The residues surrounding cluster N2, the final cluster in the FeS chain that donates electrons to ubiquinone (Figure 4.9), are clearly resolved in the active mouse complex I map (see, for example, Arg85 in Figure 4.10), defining its environment in the mammalian enzyme (Figure 4.17a). Three basic residues, Arg85, Arg105 and His190 in NDUFS2, are close to N2 and their positive charges may explain its high reduction potential (Hirst and Roessler, 2016). Arg85 is widely conserved but absent from *Thermus thermophilus*, in which N2 has a lower reduction potential (Hirst and Roessler, 2016). Furthermore, it is symmetrically dimethylated in all species tested except *E. coli* (Carroll et al., 2013), in which the N2 reduction potential is also unusually low (Hirst and Roessler, 2016). Protonation of His190 upon N2 reduction explains why N2 has a weakly pH-dependent reduction potential (Le Breton et al., 2017; Zwicker et al., 2006). Although the His190Met variant of *Yarrowia lipolytica* complex I remains catalytically competent (Le Breton et al., 2017; Zwicker et al., 2006), molecular simulations have suggested His190 deprotonation is required for electron transfer from N2 to semiquinone, and that both His190 and Arg105 (which form a cation π-stacking arrangement) contribute to the pH dependence (Gamiz-Hernandez et al., 2017). Notably, no channels from the His190 sidechain to the outside could be detected in Caver analyses (even with a 0.9 Å probe radius). Conversely, similar analyses revealed channels from both NDUFS2-Asp104 (within hydrogen-bonding distance of His190) and Glu115 (hydrogen bonded to Asp104). Unusually for cryoEM-density maps, Asp104 and Glu115 both have well-resolved sidechains, so their negative charges may be neutralized by protonation (Hryc et al., 2017). Therefore, they may act as a proton reservoir, for rapid proton delivery to His190 upon N2 reduction and/or rapid proton capture upon N2 oxidation.

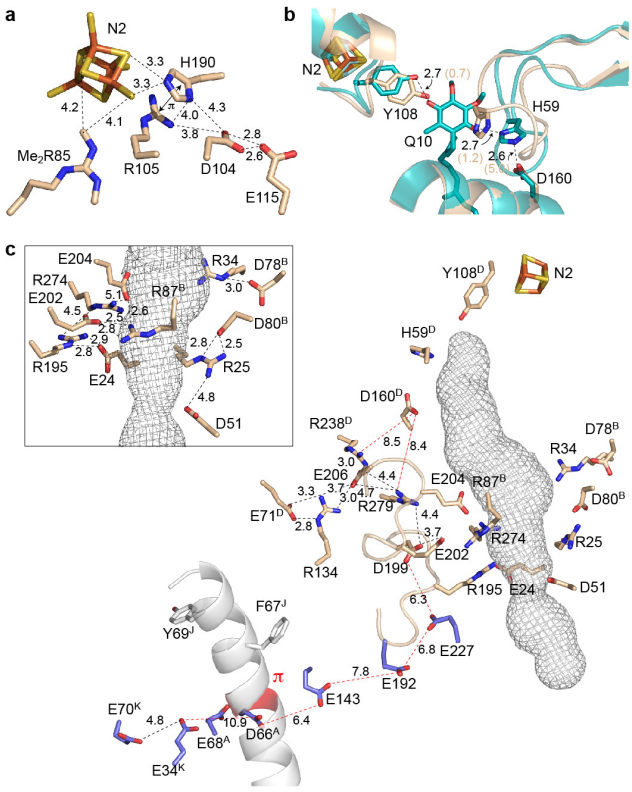


Figure 4.17: **The environment of cluster N2, the ubiquinone-binding channel and the E-channel in active mouse complex I.** (a) Cluster N2 and nearby key residues in NDUFS2. Dimethyl-Arg85 and Arg105 are positively charged, His190 is the redox-Bohr group that changes protonation state upon N2 reduction (Le Breton et al., 2017; Zwicker et al., 2006), and Asp104 and Glu115 may transfer protons to/from His190. (b) The structure around the ubiquinone-headgroup binding site with key residues in NDUFS2. The structure of active mouse complex I (teal) is overlaid on the modelled structure of bovine complex I with ubiquinone-10 bound (wheat) in which its headgroup is hydrogen bonded to the Tyr108 and His59 residues, and His59 in turn is hydrogen bonded to Asp160 (Fedor et al., 2017). Distances in brackets are for the active mouse structure. (c) The ubiquinone-binding channel (mesh) in active mouse complex I was predicted using CAVER 3.0 with a 1.4 Å probe radius. The key residues from panel b are shown with charged residues in the vicinity of the site (inset, charged residues around the centre of the channel and their interactions). The main figure highlights the carboxylate-rich ND1 TMH 5-6 loop and the E channel residues (blue) that connect the ubiquinone-binding site to the π-helix present in ND6 in the de-active enzyme, and to charged residues in the central membrane plane (Baradaran et al., 2013). Long distances of interest are in red, and all residues are in ND1 unless otherwise indicated (B - NDUFS7, D - NDUFS2, J - ND6 and K - ND4L). All distances are in Å.

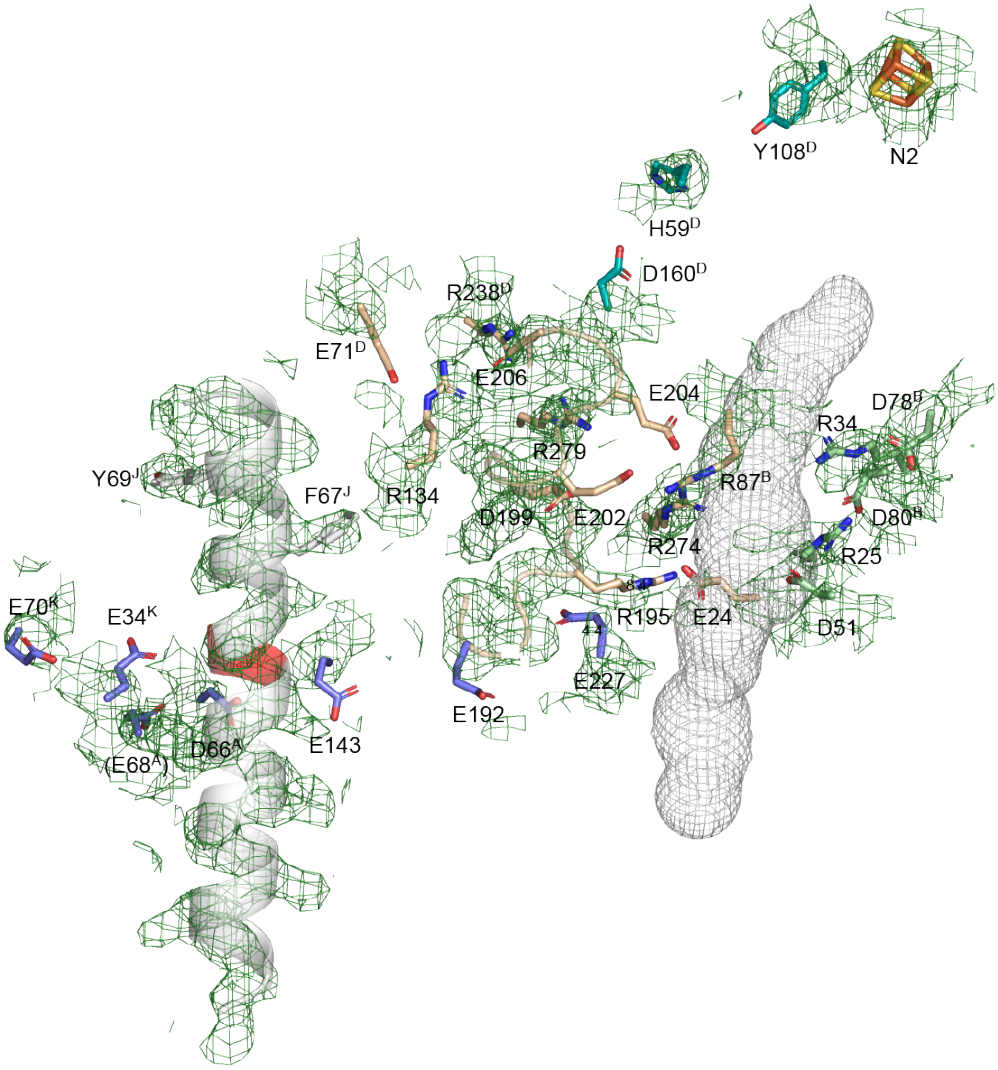


Figure 4.18: Densities around the N2 cluster and the ubiquinone site. Densities (green) were carved at 3 Å at a threshold of 0.067. The ubiquinone binding cavity is shown in grey mesh representation. Key residues are labelled.

4.10 The ubiquinone binding site and E-channel

Tyr108 and His59 in NDUFS2 are generally considered to bind the ubiquinone headgroup in complex I by forming hydrogen bonds to its redox-active C=O groups. Both residues are well resolved in the mouse map but positioned too close for ubiquinone to fit between them (5.9 Å O-to-N distance). Figure 4.17b compares the structure with a modelled structure for ubiquinone bound to bovine complex I (7.8 Å O-to-N distance) (Fedor et al., 2017), showing a substantial change in the His59 loop required for binding. Compared to previous structures, the mouse and porcine ubiquinone-binding sites match closely (Wu et al., 2016), but differ from the ovine and *Y. lipolytica* structures (Fiedorczuk et al., 2016; Zickermann et al., 2015), which have markedly different His59-loop conformations. The variation confirms the flexibility of the site, suggesting it is configured by the substrate or inhibitor binding. Finally, in the mouse structure, the His59 to NDUFS2-Asp160 (O-to-N) distance is 5.0 Å, whereas in the modelled ubiquinone-bound bovine and *T. thermophilus* structures the two are within hydrogen-bonding distance (Baradaran et al., 2013; Fedor et al., 2017). Although Asp160 has been proposed to undergo large movements upon ubiquinone reduction to initiate proton translocation (Sharma et al., 2015), its position on a nearby helix is conserved in all structures reported so far.

Figure 4.17c and 4.18 shows how networks of charged residues surround the central section (isoprenoids ~4-7) of the predicted ubiquinone-binding channel, with the regions before and after dominated by hydrophobic residues (Fedor et al., 2017). As expected for cryo-EM maps, the Arginine sidechains in Figure 4.17c are better defined than the (negatively-charged) carboxylates (Hryc et al., 2017). Nevertheless, tight interactions are observed within the networks (Figure 4.17c, inset) that, by neutralizing individual charges and restricting hydration, may assist the hydrophobic isoprenoids to bind. The charged networks may also be involved in coupling ubiquinone reduction to proton translocation. Carboxylate residues on the ND1-TMH5-6 loop (that becomes disordered in the de-active state, but exhibits similar conformations in the active mouse and bovine, porcine, human and *T. thermophilus* structures) contribute to the 'E-channel' (Baradaran et al., 2013; Guo et al., 2017; Wu et al., 2016; Zhu et al., 2016) (Figure 4.17c), which connects them (and thus the ubiquinone-binding site) to the π-bulge in ND6-TMH3 and two glutamates in ND4L, presenting a route for energy transfer to the chain of charged residues through ND2, ND4 and ND5.

4.11 π-helices in ND2, ND4 and ND5

The antiporter-like subunits ND2, ND4 and ND5 contain a common core structure of ten TMHs, TMHs 4–13 in ND4/ND5 and 1–10 in mammalian ND2, which lacks three N-terminal TMHs (Birrell and Hirst, 2010). For simplicity, the numbering system for ND4/ND5 is used throughout. TMHs 4–8 and 9–13 comprise two symmetry-related domains that face opposite sides of the membrane in a front-to-back arrangement related by rotation in an axis along the membrane domain (Efremov and Sazanov, 2011). As first observed in *E. coli* complex I (Efremov and Sazanov, 2011), symmetry-related TMHs 7 and 12 are discontinuous (interrupted in the center by substantial loops), and TMH 8, but not its symmetry-related TMH13, is distorted. The six domains of ND2, ND4 and ND5 thus form alternating 'layers' of interrupted/distorted (7-8 and 12) and continuous-α-helical (4-6 and 9-11) TMHs, along the membrane domain (Figure 4.19a).

The distortions in TMH8 of ND2, ND4 and ND5 have been described as π-bulges, although the formal ($i + 5 \rightarrow i$) hydrogen-bonding pattern is not always fulfilled in current structures, and they are clearly distorted outside the π-bulges also (Figure 4.21a and 4.22 for densities of π helices). Formal π-bulges are present in ND2-TMH8 only in the *T. thermophilus* and human respirasome structures (Baradaran et al., 2013; Guo et al., 2017), although the mouse and human structures match closely. The π-bulge is centred on Leu131, close to Lys135, a conserved charged residue in the central membrane plane (Figures 4.18c and 4.19a). Comparison of all available structures, particularly the higher-resolution (≤3.3 Å) structures from mouse, *T. thermophilus* and *E. coli* (Baradaran et al., 2013; Efremov and Sazanov, 2011), revealed no substantial differences in ND2-TMH8 or Lys135. In contrast, in ND4-TMH8, formal π-bulges are present in a common position in the *E. coli*, human and ovine structures (Efremov and Sazanov, 2011; Fiedorczuk et al., 2016; Guo et al., 2017), with the mouse structure matching its human counterpart closely, but differing in the *T. thermophilus* complex I (Baradaran et al., 2013) (Figure 4.21a). Strikingly, the two π-bulges border a central unwound region that contains conserved Lys237 in different positions, rotated around the helical axis (Figure 4.21a). In ND5-TMH8, formal π-bulges were detected centred on His248, a further conserved residue, in the mouse, porcine, *E. coli* and *T. thermophilus* structures (Baradaran et al., 2013; Efremov and Sazanov, 2011; Wu et al., 2016a) (Figure 4.21a). The position of His248 varies, although the positions of Pro242 and Phe258 are conserved to each side.

Figure 4.21b compares the positions of the conserved charged residues in ND2, ND4 and ND5, and Figure 4.19c and 4.20 shows how they form a chain along the membrane domain. In each subunit a His residue on the loop in interrupted-TMH7 points outside the subunit (Figure 4.19c, grey labels) [ND4-His213 interacts with ND5-Asp554 on the transverse helix, clasping it against ND4]. Glutamate residues on TMH5 may be stabilized electrostatically by charged residues on interrupted-TMH7 and -TMH12 (Figure 4.19c, red labels). ND2-Lys135 and ND4-Lys237 (Figure 4.19c, blue labels) are on TMH8 but not conserved in ND5. Instead, ND5-His248 on TMH8 is positioned, like ND4-His220, closer to the matrix side of the membrane, suggesting a role in proton uptake. Further residues, not conserved in all three subunits, contribute additional stepping stones along the domain (Figure 4.19c, black labels).

Finally, the ND5 transverse helix, which runs alongside ND5, ND4 and ND2 (Figure 4.9), has been suggested to form an important coupling element for energy transfer (Efremov and Sazanov, 2011). It contains a substantial stretch of π-helix in the mouse structure (Figure 4.21b), and also in the porcine and *E. coli* structures (in which the C-termini of ND5 are clearly resolved) (Efremov and Sazanov, 2011; Wu et al., 2016). However, in *T. thermophilus* complex I (Baradaran et al., 2013), except for a short distortion at the start, the transverse helix is α-helical throughout. Despite these differences, the position of Asp554 is conserved and the whole transverse helix extends the same distance in each case — due to its different structures it contains three more residues in mouse than *T. thermophilus* and five more than in *E. coli*.

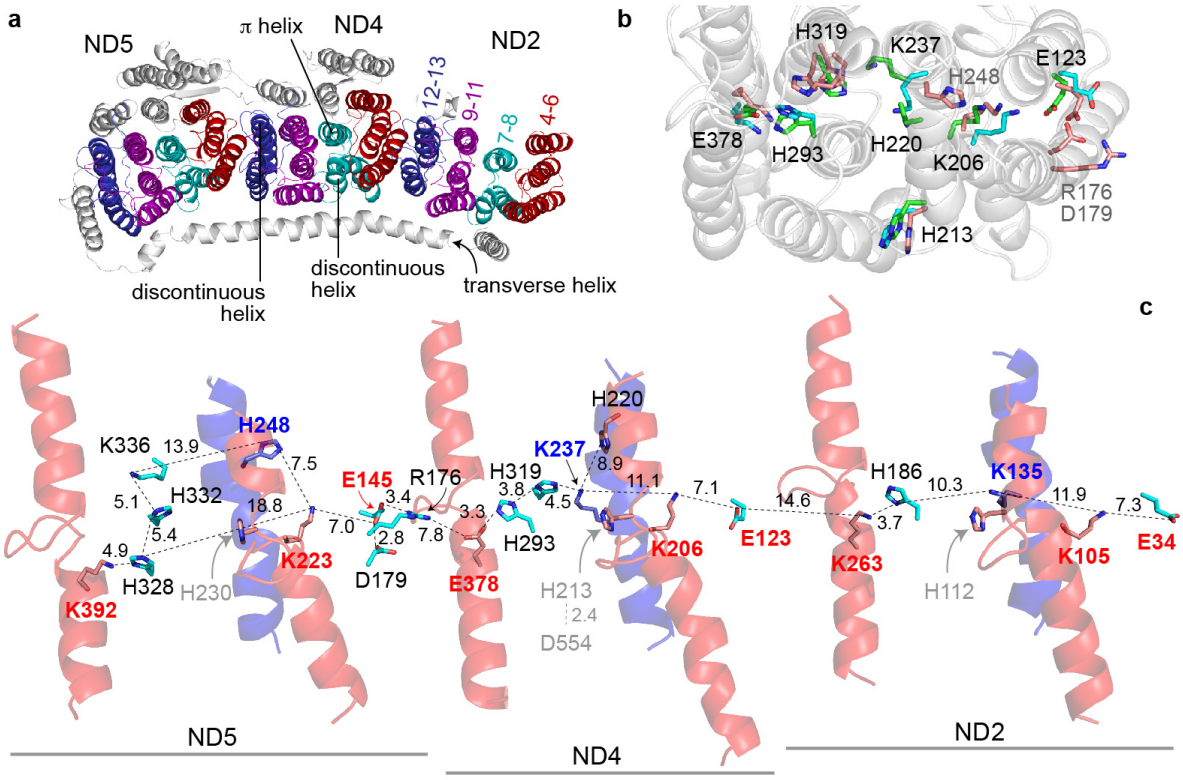


Figure 4.19: Charged residues and discontinuous and π-helices in subunits ND2, ND4 and ND5. (a) Overview of subunits ND2, ND4 and ND5 showing the layered arrangement of core TMHs 4 to 13. TMHs 4 to 6 (red) and 9 to 11 (magenta) are α helical whereas TMH7 is discontinuous and TMH8 contains π-helical structure (cyan) and TMH12 (blue) is also discontinuous. The transverse helix from ND5 is also shown. (b) Overlay of conserved charged residues in the central membrane plane in ND2, ND4 and ND5. A cartoon representation of ND4 is shown for reference, and conserved residues are coloured cyan for ND2, green for ND4 and pink for ND5. Numbers refer to ND4, except for residues only in ND5 numbered in grey. (c) The distribution of charged residues and helical elements along the membrane arm with key residues labelled and the distances between ionizable groups shown in Å. TMH7 and 12 are in red and TMH8 is in blue.

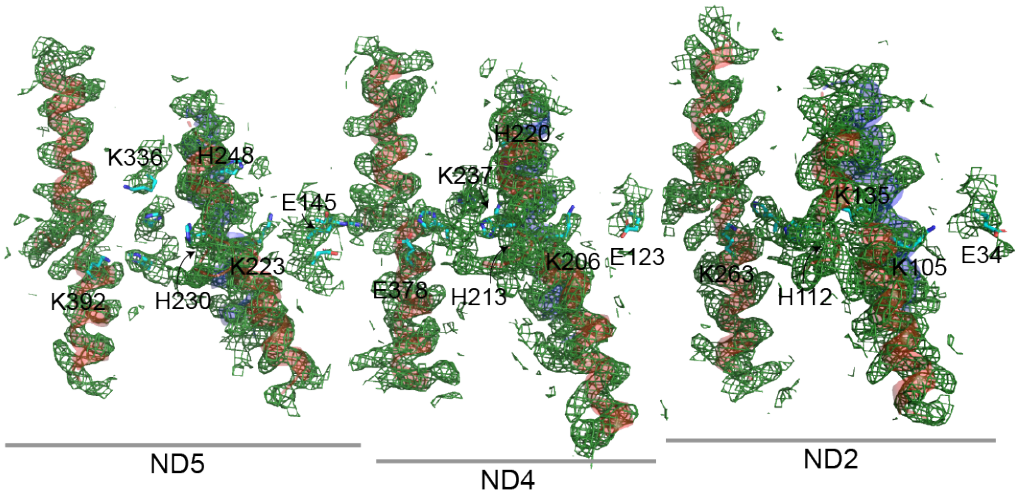


Figure 4.20: Densities of discontinuous and π-helices in subunits ND2, ND4 and ND5. Densities were carved at 3 Å at a threshold of 0.067. key residues are labelled. TMH7 and TM12 are in red and TMH8 is in blue.

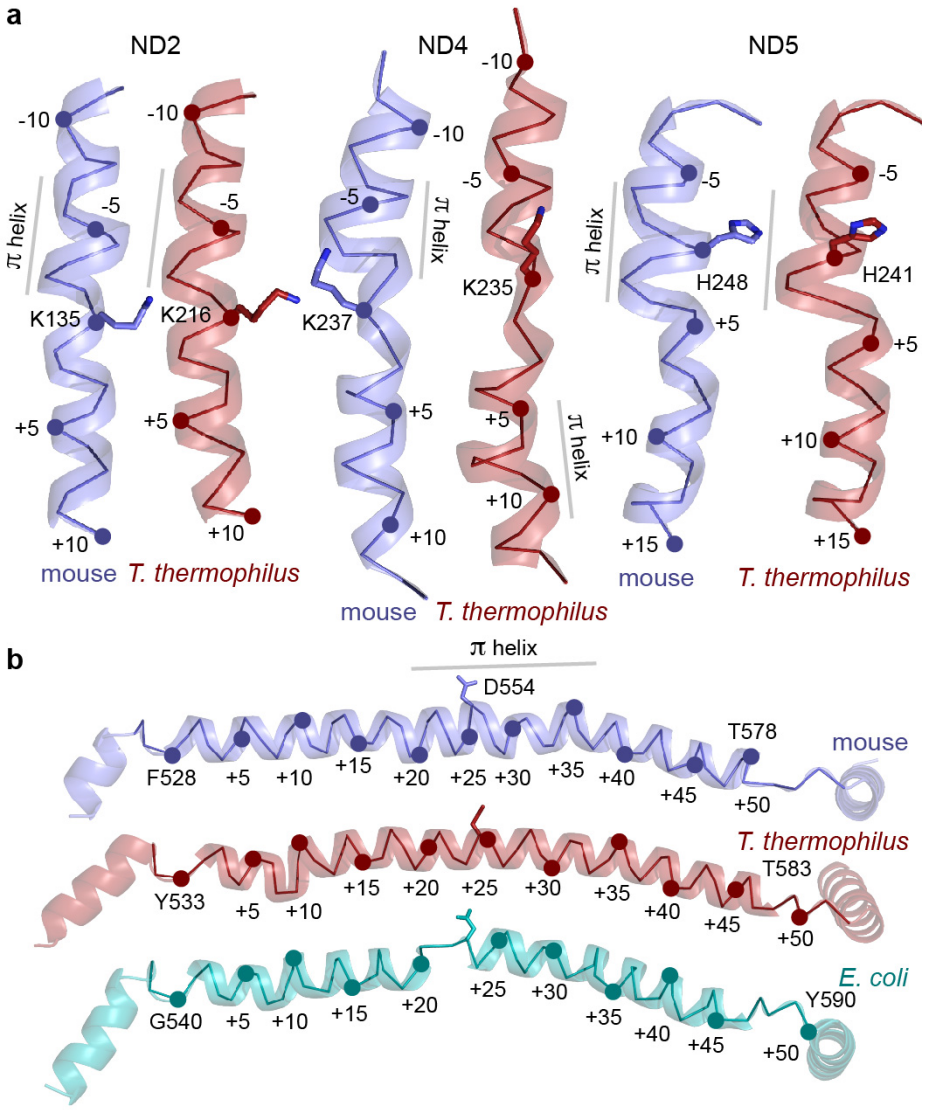


Figure 4.21 Comparison of π-helices in subunits ND2, ND4 and ND5 in mammalian and bacterial complex I structures. (a) TMH8 in ND2, ND4 and ND5 orientated with the mitochondrial matrix at the top. (b) Comparison of structures of the transverse helix in ND5 with TMH16 (the C-terminus) on the right and the conserved aspartate residue indicated. Groups of structures have been aligned using the whole subunits structure. The ribbons join the Cα atoms together, and each fifth residue is numbered. The approximate positions of stretches of π-helix are indicated. The mouse structure is for the active state, the structure of *T. thermophilus* complex I is from 4HEA.PDB (Baradaran et al., 2013) and the structure of the membrane domain of complex I from *E. coli* is from 3RKO.PDB (Efremov and Sazanov, 2011).

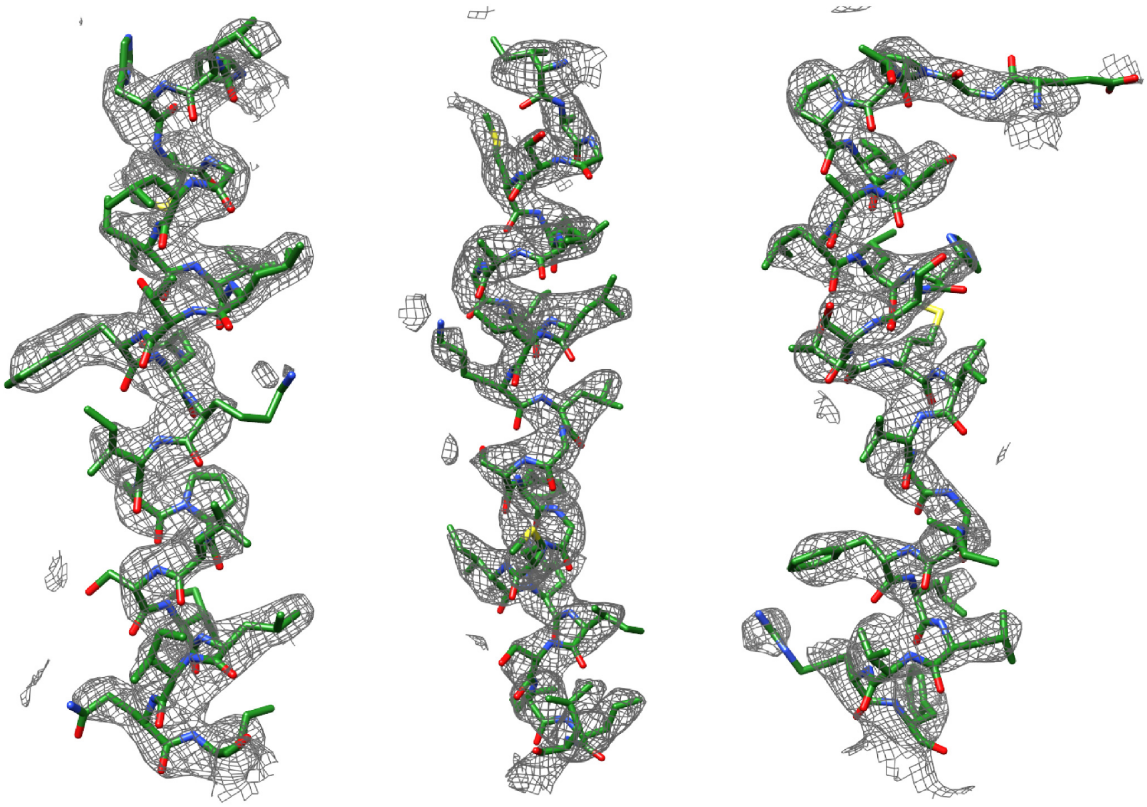


Figure 4.22: Densities of TMH8 in the active mouse complex I structure. Densities were carved at 2 Å radius with threshold level of 0.065.

4.12 Discussion

4.12.1 Resolution-based limitation of previous mammalian structures

The architecture of mammalian complex I was first resolved in bovine, and it proved crucial in determining the arrangement of some of the supernumerary subunits (Vinothkumar et al., 2014). The rest of the supernumerary subunits were later assigned when the structures of the bovine and ovine complex were resolved (Fiedorczuk et al., 2016; Zhu et al., 2016). In addition, the structure of the respiratory supercomplexes (the higher-order organisation of respiratory enzymes) were also resolved (Gu et al., 2016; Guo et al., 2017; Letts et al., 2016a; Wu et al., 2016).

However, with the exception of *T. thermophilus* and *E. coli* complex I (Baradaran et al., 2013; Efremov and Sazanov, 2011), the medium resolution structures of mammalian complex I have hampered structural analysis into key elements that may be involved in catalytic turnover (Fiedorczuk et al., 2016; Guo et al., 2017; Letts et al., 2016a; Wu et al., 2016; Zhu et al., 2016), or pathophysiological states, such as that observed in ischemia (Chouchani et al., 2014).

4.12.2 Conclusion

The structure of mouse complex I at 3.3 Å resolution, now matching the *T. thermophilus* structure in terms of resolution, has many identifiable features that was only previously inferred from mass spectrometry (Carroll et al., 2009), such as the symmetrically demethylated modification of Arg85 (see Figure 4.10). This arginine modification is important as it is suggested to modulate the midpoint potential of the terminal iron-sulphur cluster of complex I, N2 (Hirst and Roessler, 2016).

In addition, a novel density detected in the active site of the nucleoside kinase subunit of complex I, NDUFA10, has been modelled as ADP. Comparison with bovine complex I show that the nucleotide binding site is precluded by a phosphorylated serine not conserved in the mouse enzyme. Despite the nucleotide being bound to the catalytic site of NDUFA10, its bound in an inhibitive state, resembling the feedback inhibition state exhibited in the deoxyribonucleoside kinase of *Drosophila melanogaster* (Figure 4.15) (Mikkelsen et al., 2003).

Finally, this study has shown that transmembrane helices (TMH) of the mitochondrially encoded subunits are able to adopt different confirmations/distortions within the same species and between species. For example, TMH3 of ND6 can forms a π-helix from an α-helix upon induced de-activation of the mouse enzyme. The π-helix forming TMH3 of ND6 connects the ubiquinone binding cavity to the charged residues lining the membrane domain, and therefore under catalysis, this may potentially be important in coupling redox reaction at the Q-site to proton translocation in the membrane domain. TMH8 in ND2, ND5 and ND4 also form π-helices, in particular, the ND4 π-helix adopts a different conformation in mouse complex I compared to the *T. thermophilus* enzyme, which may be crucial for proton uptake, as this helix is closer to the matrix/cytoplasmic side.

4.12.2 Future prospects

Even at 3.3 Å, the resolution is still limited as water molecules are not clearly resolved and therefore cannot be assigned in the structure. Water molecules are essential in proton conductance in complex I; however, their presence has only been predicted but never visualised (Sharma et al., 2015). Therefore in order to improve the resolution of complex I, the purified mouse enzyme must be imaged at an even higher magnification than what is used in this study, and using the K2 summit camera, as this would improve the detective quantum efficiency at higher special frequency (higher resolution) (McMullan et al., 2016).

As all the structures of complex I are in the oxidised state, resolving the structure in the reduced state would prove crucial in understanding the catalytic cycle of the enzyme and how key elements in the structure may change upon reduction.

Chapter 5: Molecular characterisation of complex I from the *Ndufs4 (-/-)* mouse strain

5.1 Introduction

5.1.1 The NDUFS4 subunit of complex I

NDUFS4 (18 kDa in bovine nomenclature) is a nuclear-encoded supernumerary subunit of complex I located in the matrix arm of the enzyme (Figure 5.1). The *Ndufs4* gene in mouse is made up of 5 exons, encoding a protein of 133 amino acids. NDUFS4 is a highly conserved protein and one of three supernumerary subunits (along with NDUFA12 and NDUFS6), that are conserved between mammals and the mitochondria-related alphaproteobacterium *Paracoccus denitrificans* (Yip et al., 2011).

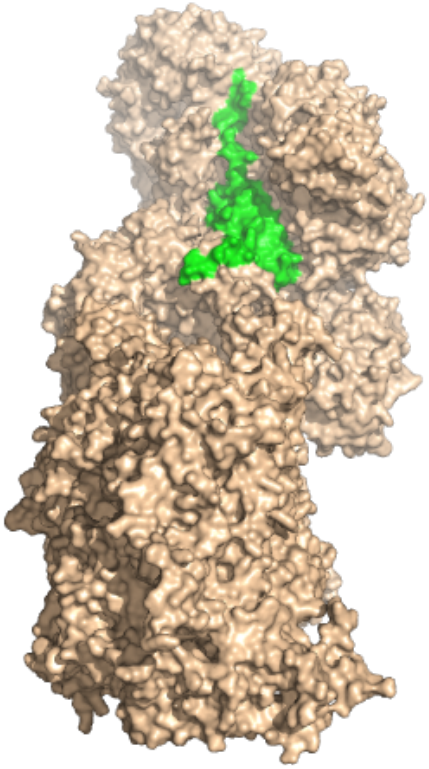


Figure 5.1: Position of NDUFS4 on complex I. Surface representation of NDUFS4 (green) on complex I (wheat). The structure is of the active mouse complex I presented in Chapter 4.

Early reports of NDUFS4 in bovine heart mitochondria suggested that it was phosphorylated at serine 131 by cAMP-dependent protein kinase (Papa et al., 1996). Others further reported that phosphorylation of NDUFS4 in mouse fibroblasts increased rotenone-sensitive oxygen consumption (Scacco et al., 2000), thus proposing NDUFS4 phosphorylation has a regulatory role for the catalysis of complex I. However, in-depth proteomic analyses of cAMP phosphorylation of complex I showed that the phosphorylation site was in NDUFB11 and not NDUFS4. The confusion arose because of their similar masses, leading to miss-identification of the phosphorylated protein. In spite of this, loss of regulation of NDUFS4 by phosphorylation is still mistaken as the cause of complex I deficiencies when the said subunit is deleted (Anderson et al., 2008; van den Heuvel et al., 1998).

5.1.2 Mutations in *Ndufs4* and associated clinical phenotypes

Pathological mutations in *Ndufs4* result in complex I deficiency in humans, and animal models (Anderson et al., 2008; Kruse et al., 2008; Petruzzella et al., 2001; Scacco et al., 2003). Pathological *Ndufs4* mutations result in the loss of protein but are always autosomal recessive, meaning both genes have to be affected by a severe phenotype to manifest, resulting in a functional knockout of the protein. Patients with these mutations display Leigh syndrome, which is an early onset neurodegenerative encephalopathy. Presentation of this disease includes symmetrical bilateral lesions in the basal ganglia and other parts of the brain, visible by magnetic resonance imaging (Figure 5.2) (Ortigoza-Escobar et al., 2016). Clinical symptoms include, but are not limited to, lactic acid accumulation in blood plasma, lethargy, ataxia, loss of motor skills, and as the disease progresses, death (Budde et al., 2003; Ortigoza-Escobar et al., 2016). 22 reported clinical cases of Ndufs4 mutations have been reviewed in Ortigoza *et al*, 2016 (Ortigoza-Escobar et al., 2016).

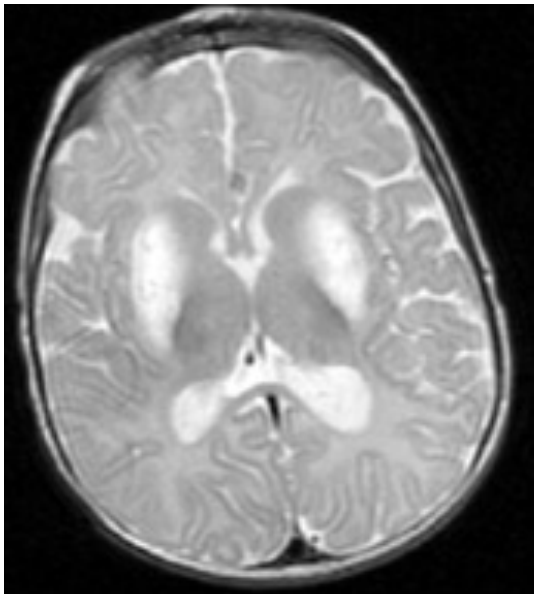


Figure 5.2: Magnetic resonance imaging of a cross-section of the brain of a patient with a premature stop codon in the *Ndufs4* gene (Trp97Ter) display white lesions in the basal ganglia. This figure was taken from Ortigoza *et al*, 2016 (Ortigoza-Escobar et al., 2016).

5.1.3 The *Y. lipolytica Ndufs4* deletion strain

The NDUFS4 (NUYM) subunit in the aerobic yeast, *Y. lipolytica* was removed by deleting the entire open reading frame encoding the protein, using homologous recombination (Kahlhöfer et al., 2017). As a result of the mutation, complex I content and catalytic activity (NADH:DQ oxidoreduction) in mitochondrial membranes is significantly diminished when compared to its wild type counterpart. Interestingly, complex I from the knockout strain appears intact on BN-PAGE. This is in contrast to human and mouse models that lack NDUFS4, in which subassemblies of complex I are observed alongside the intact enzyme (Calvaruso et al., 2012; Ugalde et al., 2004). Electrophoretic analyses of purified complex I from the mutant strain revealed the presence of NDUFAF2 (N7BML), which is a putative assembly factor of complex I (Kahlhöfer et al., 2017). However, this assembly factor was detected at a substoichiometric level in the purified sample, making its role unclear in the mutant strain. Comparison of the EPR spectra of the iron-sulphur clusters between the wild type and knockout strain, showed that the absence of NDUFS4 results in their decreased signal, in particular, N1b. This means that the clusters of complex I are not as readily reduced in the knockout strain when compared to its wild type counterpart. Surprisingly, the flavin site activity (using HAR) of both strains were essentially indistinguishable despite a reduction in NADH:DQ activity, suggesting a

problem with intramolecular electron transfer. Although the study provided plenty of information regarding how NDUFS4 deletion affects complex I content and activity, the yeast model system is limited and cannot fully recapitulate the effects seen in mammals (Kruse et al., 2008).

5.1.4 A mouse model with an *Ndufs4* deletion

In order to generate a mouse model based on the human mutations of *Ndufs4*, the second exon of *Ndufs4* was removed by Cre Recombinase-mediated homologous recombination (using exon 2 flanking loxP sites), resulting in a gene only coding for the first 14 amino acid residues of the mitochondrial targeting sequence (Kruse et al., 2008). NDUFS4 could not be detected by Western blotting when both copies of the gene were affected, therefore the mouse model is essentially an NDUFS4 knockout or an *Ndufs4* ($^{-/-}$), (which is similar to humans with mutations in *NDUFS4*). Mice with the mutation presented signs of neurodegeneration, including lethargy, ataxia, and loss of motor skills. Similar to humans, the progression of the disease resulted in early death compared to wild type mice (~50 days). The *Ndufs4* ($^{-/-}$) mice model is able to broadly recapitulate the effects seen in humans with mutations in *Ndufs4*.

Complex I from mice with NDUFS4 absent had a reduced NADH:DQ activity and content in mitochondria, as with human patients with mutations in the gene (Antonicka et al., 2003; Kruse et al., 2008; Ugalde et al., 2004). Moreover, there have been some reports of the visualisation of a stalled complex I assembly intermediate (~830 kDa) by BN-PAGE in the *Ndufs4* ($^{-/-}$) mouse model (Calvaruso et al., 2012; Valsecchi et al., 2012), also observed in humans (Ugalde et al., 2004). As this observed putative assembly intermediate did not contain the flavoprotein domain (site of NADH oxidation). It was reasoned that NDUFS4 may be required for the assembly of the enzyme (Calvaruso et al., 2012).

5.1.5 Deletion of *Ndufs4* in *Arabidopsis thaliana*

As with mice and aerobic fungi, the *Ndufs4* gene was also deleted in *A. thaliana* to observe the effect it would have on complex I assembly (Meyer et al., 2011). In the *Ndufs4*-deleted plant strain, there was no detectable flavin-sensitive NBT staining by BN-PAGE analysis, meaning no mature complex I was present for that strain. This is dissimilar to the *Ndufs4* ($^{-/-}$) fungus and mice strain where although complex I content is reduced, the intact enzyme could still be

detected by BN-PAGE analysis (Calvaruso et al., 2012; Kahlhöfer et al., 2017). Moreover the sub-assembly of complex I contained mostly proteins from the membrane arm, indicating that its assembly may differ from that of the mouse enzyme, where the corresponding deletion results in the detection of a larger subcomplex containing both the membrane arm and lower half of the matrix arm (Calvaruso et al., 2012).

5.1.6 Potential therapies

Ndufs4 ($^{-/-}$) mice treated with rapamycin, an inhibitor of the mechanistic target of rapamycin (mTOR), showed improved longevity, and an ameliorated neurological disorder compared to the untreated *Ndufs4* ($^{-/-}$) mice (Johnson et al., 2013). It was suggested that a rewiring of metabolism resulted in a decrease in glycolysis and catabolism, preventing the build of excess glycolytic intermediates in mice. However, treatment with rapamycin retarded the growth of the mice and in addition, because of the short half-life of rapamycin in the blood meant that daily treatment with the drug was required.

A genome-wide editing study (Cas9 mediated) on cell lines treated with antimycin A (an inhibitor of complex III) showed that deletion of von Hippel-Lindau (VHL), a ubiquitin ligase, is protective against defects in the respiratory chain by an increase in cell growth compared to control cells (Jain et al., 2016). Under normoxia (~20% oxygen), the hypoxia-inducible transcription factors (HIFs), which are responsible for the hypoxic response, are constitutively made and marked for degradation by VHL, but under hypoxia, HIFs are stabilised and are transcriptionally active resulting in the up-regulation of hypoxia-dependent changes (Choudhry and Harris, 2018). To investigate whether there was a link between hypoxia and protection against respiratory defects, *Ndufs4* ($^{-/-}$) mice were subjected to chronic hypoxia (~11% oxygen) and as a result, mice survived for longer, displayed normal growth and showed few neurological symptoms, making it phenotypically indistinguishable from wild type mice (Jain et al., 2016). Later work even showed a reversal of brain lesions after hypoxic exposure (Ferrari et al., 2017). Attempts to inhibit VHL in mice in order to initiate the hypoxic response (to alleviate some of the symptoms of the Leigh-like syndrome) were unsuccessful because of obstruction of drug delivery to the central nervous system by the blood-brain barrier (Jain et al., 2016). Although hypoxia therapy appears effective in alleviating neurological symptoms for this mouse model,

application of this treatment must be tried in other mitochondrial disease models that affect respiration before trials on humans can begin.

Logically, the most reasonable treatment for the *Ndufs4* deletion is to re-introduce the missing subunit to all tissues. Introducing a viral vector carrying the human version of *Ndufs4* gene into *Ndufs4* ($^{-/-}$) mice, restored complex I in liver mitochondria and partially alleviated the clinical phenotype (Di Meo et al., 2017). The main problem encountered in this study was the uniform delivery of the viral vectors to all tissues. In particular, the blood-brain barrier obstructs the efficient delivery of the vector to the brain, the most severely affected organ.

5.2 Aims of this chapter

The work in this chapter aimed to:

- Using a proteomic approach, establish whether the absence of NDUFS4 affects the maturation of complex I.
- Characterise complex I from the *Ndufs4* ($^{-/-}$) mice, in order to understand the underlying mechanism of its dysfunction.

5.3 Genotyping *Ndufs4* mice

The severity of *Ndufs4* ($^{-/-}$) mice phenotype makes breeding from ($^{-/-}$) breeding pairs unfeasible, therefore, in order to maintain a progeny with the *Ndufs4* disruption, a heterozygous breeding pair is required *Ndufs4* ($^{+/-}$). This theoretically results in a progeny that is 50% ($^{+/-}$), 25% ($^{+/+}$) and 25% ($^{-/-}$) for NDUFS4, according to Mendelian genetics. Prior to any characterisation of complex I from mice, the status of disruption of *Ndufs4* from the genome of mouse pups had to be confirmed. Genomic DNA was extracted from mouse ear biopsy tissue using a Maxwell® 16 DNA Purification Kit according to the manufacturer's protocol. Thereafter, DNA regions were amplified using PCR, with the DNA primers used and the conditions for the reaction described in detail in Chapter 2.2. The resulting products were analysed by agarose gel electrophoresis (Figure 5.3). The presence of *Ndufs4* exon 2 was confirmed by amplification of a genomic region, resulting in a single band ~150 base pairs in size (Figure 5.3a). Similarly, disruption of *Ndufs4* exon 2 was confirmed by amplification of a product that is ~200-300 base pairs in size (Figure 5.3b). A simplified scheme of both reactions is shown in figure 5.3c. Based on complementary information from both reactions, mice were genotyped heterozygous *Ndufs4* ($^{+/-}$), homozygous *Ndufs4* ($^{-/-}$) or wild type *Ndufs4* ($^{+/+}$). Furthermore, the genotyping was confirmed for *Ndufs4* ($^{-/-}$) mice, as they display a severe phenotype (growth and mobility defects) compared to wild type *Ndufs4* ($^{+/+}$) and heterozygous *Ndufs4* ($^{+/-}$) mice, which displayed healthy phenotypes (Kruse et al., 2008).

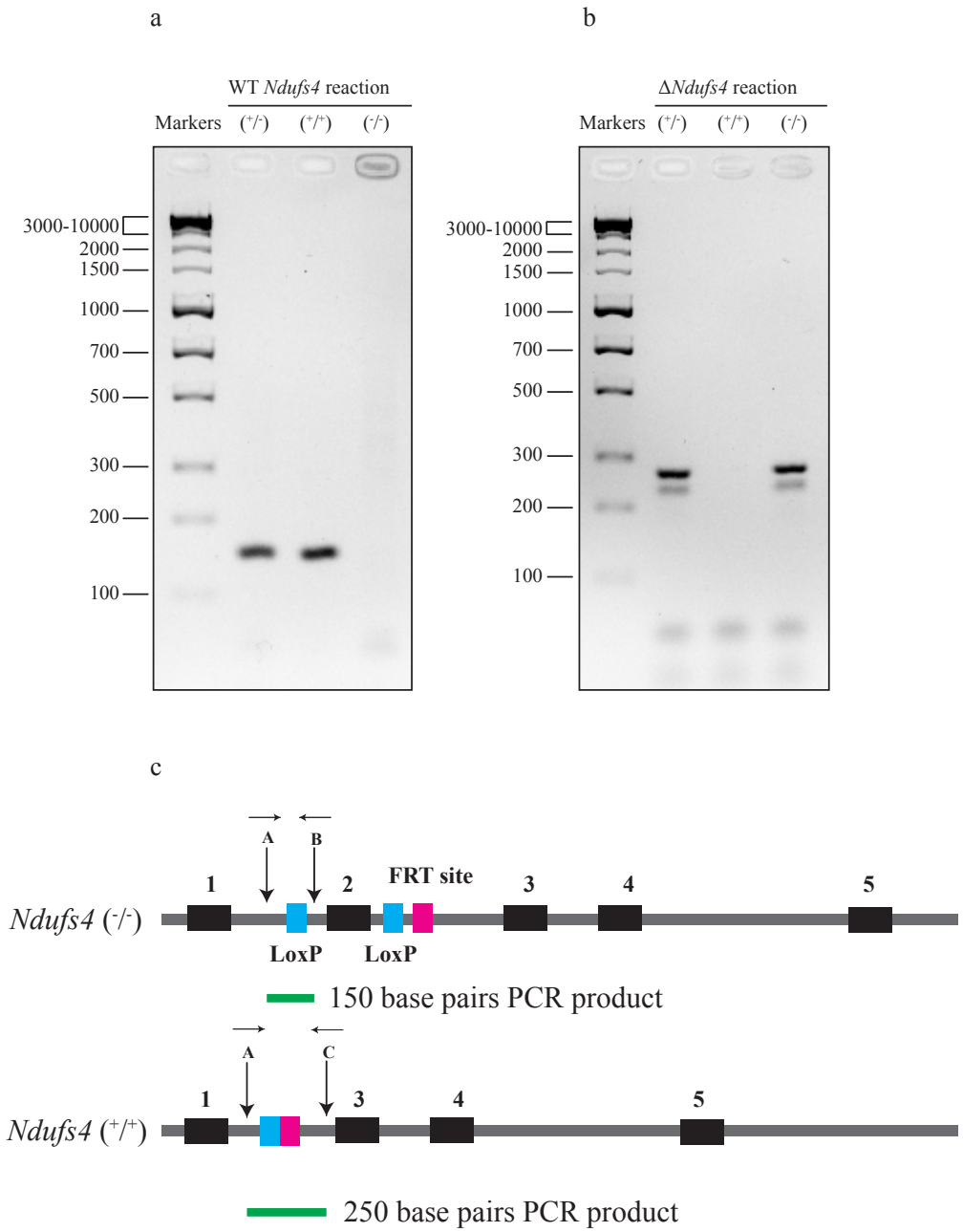


Figure 5.3: Genotyping mice (ear biopsy) for *Ndufs4* deletion. PCR and electrophoresis were carried out as described in Chapter 2.2**.** PCR products (5 µl) were separated by agarose gel electrophoresis, then DNA products were stained with ethidium bromide and visualised using a Biorad imaging system. Known DNA markers (in base pairs) were loaded for reference. The same genomic DNA was used for both PCR, to confirm the presence (a) or absence (b) of *Ndufs4* exon 2. (c) A scheme of both reactions resulting in the genotyping mice as heterozygous *Ndufs4* ($^{+}/^{-}$), homozygous *Ndufs4* ($^{-}/^{-}$) or homozygous *Ndufs4* ($^{+}/^{+}$). Exons, LoxP sites, the FRT site and PCR products are shown as black, blue, pink and green, respectively.

5.4 Analyses of *Ndufs4* ($^{-/-}$) membranes

First mitochondria and then membranes were prepared from *Ndufs4* ($^{-/-}$) heart as described for *Ndufs4* ($^{+/+}$) in Chapter 2.4. As *Ndufs4* ($^{-/-}$) mice suffer from complex I deficiency, with an associated decrease in mitochondrial quality compared to *Ndufs4* ($^{+/+}$) (Kruse et al., 2008), solubilised membrane protein (1% DDM) from both strains were compared by BN-PAGE (Figure 5.4). In agreement with previous studies, intact complex I was significantly reduced in *Ndufs4* ($^{-/-}$) membranes compared to the *Ndufs4* ($^{+/+}$) membranes, while the amount of the other oxidative phosphorylation enzymes appear unchanged on BN-PAGE.

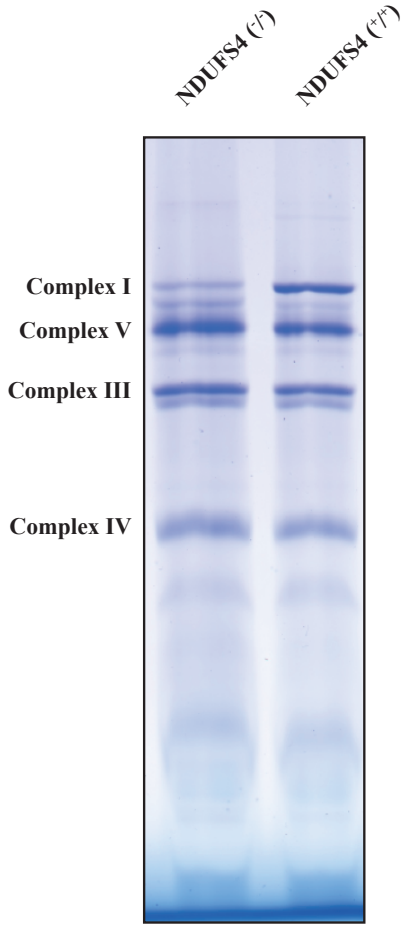


Figure 5.4: Comparative BN-PAGE of *Ndufs4* ($^{-/-}$) and *Ndufs4* ($^{+/+}$) membranes. Membranes (10 µg) were solubilised in 1% DDM and analysed by BN-PAGE as described in Chapter 2.5.1.

Catalytic activities were also measured for *Ndufs4* ($^{-/-}$) mouse membranes and compared to *Ndufs4* ($^{+/+}$) mouse membranes (as described in Chapter 2.6). Complex I-dependent NADH:O_2 measurements (catalysis by complexes I, III and IV) of *Ndufs4* ($^{-/-}$) membranes showed only 10% of the catalytic activity of *Ndufs4* ($^{+/+}$) membranes, as expected for a complex I-deficient strain. This result is consistent with the decrease in complex I observed by BN-PAGE (Figure 5.4). Complex I-independent succinate:O_2 rates (catalysis by complexes II, III and IV) were essentially the same for *Ndufs4* ($^{-/-}$) membranes and *Ndufs4* ($^{+/+}$) membranes, confirming that, for the knockout, complex I is rate limiting for catalysis (Table 5.1). However, relative NADH oxidation activities that measure the rate of flavin-site reactions of *Ndufs4* ($^{-/-}$) membranes compared to *Ndufs4* ($^{+/+}$) membranes varied substantially with the oxidising substrate used (FeCN, HAR and APAD). The relative activities of *Ndufs4* ($^{-/-}$) membranes to *Ndufs4* ($^{+/+}$) were 6% with HAR, 16% with FeCN and 25% with APAD (Table 5.1). Although the flavin-site assays are in agreement with the result from BN-PAGE (Figure 5.4), in which complex I is markedly reduced in *Ndufs4* ($^{-/-}$) membranes, the inconsistency in the relative activities, coupled with the different mode of action by the artificial substrates used (Birrell et al., 2011, 2009), suggests that in addition to reduced abundance, there is a change in the flavin-site characteristics of the enzyme as a result of the subunit deletion.

As complex I is the main source of reactive oxygen species in mitochondrial membranes (Pryde and Hirst, 2011), mutations in complex I are often correlated with increased superoxide production. H_2O_2 production in membranes, which is a product of the rapid dismutation of superoxide, was measured by monitoring the oxidation of Amplex Red to resorufin (Table 5.1). Interestingly, H_2O_2 production by *Ndufs4* ($^{-/-}$) membranes was 65% of that of *Ndufs4* ($^{+/+}$) membranes. H_2O_2 is also produced by the flavin-site, so this is inconsistent with the range of relative activities (6-25%) observed in the other flavin-site reactions (Table 5.1). Considering membranes contain a complicated mixture of proteins, not all of the reactive oxygen species produced may be attributed to complex I, particularly in the case of *Ndufs4* ($^{-/-}$) membranes, with reduced complex I content. Alternatively, there could also be a change in the flavin-site characteristics as a result of the subunit deletion.

Finally, hydrolysis of ATP by F_1F_O-ATPase was measured in membranes, and it was observed that the activity of *Ndufs4* ($^{-/-}$) membranes is noticeably lower than *Ndufs4* ($^{+/+}$) membranes (68% relative activity), for reasons that are unclear (Table 5.1).

Measurement (μmol min^{-1} mg^{-1})	*Ndufs4* ($^{-/-}$)	*Ndufs4* ($^{+/+}$)	Relative activities (%)
NADH:O_2	0.15 ± 0.002	1.47 ± 0.08	10
Succinate:O_2	1.18 ± 0.04	1.22 ± 0.02	97
ATP hydrolysis	2.58 ± 0.02	3.82 ± 0.07	68
H_2O_2	$2.28\times10^{-3} \pm 0.06\times10^{-3}$	$3.53\times10^{-3} \pm 0.12\times10^{-3}$	65
FeCN	0.83 ± 0.03	5.13 ± 0.07	16
HAR	0.35 ± 0.02	5.71 ± 0.06	6
APAD	0.19 ± 0.006	0.75 ± 0.02	25

Table 5.1: Enzymatic activity measurements on *Ndufs4* ($^{-/-}$) and *Ndufs4* ($^{+/+}$) mouse membranes. Enzymatic activity measurements were carried as described in chapter 2.6. The rates presented for *Ndufs4* ($^{+/+}$) membranes are taken from Table 3.2. The data are presented as mean values with standard error of the mean (n ≥ 3).

The catalytic activities measured in Table 5.1 in conjunction with BN-PAGE evidence suggest that both complex I amount and catalytic function is severely affected as a result of the subunit deletion. To investigate whether the altered flavin-site reactions in *Ndufs4* ($^{-/-}$) membranes result from a structural rearrangement of the enzyme (for example stalled assembly intermediates), membranes from both strains were analysed and compared using the proteomic BN-PAGE complexome profiling technique as described in detail in Chapter 2.9.6 (Figure 5.5) (Heide et al., 2012). In brief, membranes were solubilised by 1% DDM, then the solubilised proteins were separated by BN-PAGE and stained with Coomassie Blue. Thereafter, the gel was excised into 64 equal-sized slices for tryptic digestion, and the resulting peptides were sequentially analysed by LC-MS.

In agreement with previous analysis (Figure 5.4), the BN-PAGE showed a reduced complex I content in *Ndufs4* ($^{-/-}$) membranes compared to *Ndufs4* ($^{+/+}$) membranes, but similar amounts for the other oxidative phosphorylation enzymes (Figure 5.5a, b).

To gain insight into new interactions between complex I and other proteins as a result of the knockout, hierarchical clustering of proteins was carried out, based on migration pattern on the gel (Heide et al., 2012). In essence, the three top scoring peptides for each protein from the LC-MS analyses were chosen as representative peptides for the same protein. For proteins detected

in the gel slices, normalised scores were assigned based on their maximum scores along the gel, visualised with the corresponding heatmap (Figure 5.5a, b). The heatmap was filtered to highlight credible protein interactions with complex I (i.e. known assembly factors) in *Ndufs4* ($^{-/-}$) membranes, using the following criteria:

- A similar migration pattern of proteins to known complex I subunits.
- A change in the migration pattern of proteins along the gel in *Ndufs4* ($^{-/-}$) membranes compared to *Ndufs4* ($^{+/+}$) membranes, signifying a functional change in proteins not affected by gel electrophoresis.
- An apparent increase in the amount of proteins in *Ndufs4* ($^{-/-}$) membranes compared to *Ndufs4* ($^{+/+}$) membranes, indicating their importance in the knockout.

As observed in the heatmaps, 41 out of 44, and 39 out of 43 (minus NDUFS4) unique subunits of complex I were detected in the analyses of *Ndufs4* ($^{+/+}$) membranes and *Ndufs4* ($^{-/-}$) membranes, respectively (Figure 5.5a, b). ND6, ND3 and ND4L were not detected in either membranes (evident by their absence in both heatmaps), owing to the few tryptic cleavage sites present in those subunits. In the heatmaps, NDUFB5 shows aberrant migration patterns from the rest of the complex I subunits by appearing at a higher molecular position on the gel, in both *Ndufs4* ($^{-/-}$) and *Ndufs4* ($^{+/+}$) membranes, suggestive of a BN-PAGE artefact. ND1 appears to only migrate with the complex I cluster (slice 23 in heatmap) in *Ndufs4* ($^{-/-}$) membranes, but not with complex I (slice 21 in heatmap) in *Ndufs4* ($^{+/+}$) membranes (Figure 5.5a, b). However, because of the poor tryptic peptides produced from ND1, its detection is likely masked by complex I subunits rich in tryptic peptides (especially in *Ndufs4* ($^{+/+}$) membranes where complex I content is high). NDUFC1 was not detected in *Ndufs4* ($^{-/-}$) membranes, unlike in *Ndufs4* ($^{+/+}$) membranes. NDUFC1 has few tryptic peptides, making it difficult to detect, particularly if the complex I content is reduced.

As observed on the heatmaps, there are three significant complex I clusters in *Ndufs4* ($^{+/+}$) membranes and *Ndufs4* ($^{-/-}$) membranes, respectively (Figure 5.5). The highest of those complex I clusters on the gel correspond to intact complex I for both samples. The intact complex I cluster is the least abundant in *Ndufs4* ($^{-/-}$) membranes, as observed by its weak intensity on the heatmap (Figure 5.5b). For *Ndufs4* ($^{-/-}$) membranes the second cluster (main cluster) appears in the same position as complex V on the gel (slice 24 on the heatmap); it is

termed the 830 kDa complex following previous nomenclature (Ogilvie et al., 2005), and contains 30 of the 44 unique subunits of complex I. The 830 kDa complex comprises the membrane domain and most of the quinone domain (Q-domain), with the exception of NDUFA7, which is mostly detected in slice 58 on the heatmap (Figure 5.5b). According to the heatmap, the 830 kDa complex also lacks the flavin domain (N-domain), therefore it cannot carry out NADH oxidation, making it undetectable with NBT staining (see Figure 5.9). Lastly, the smallest complex I cluster of *Ndufs4* ($^{-/-}$) membranes, which happens to co-migrate with complex IV on the gel (slice 37) is termed the flavin complex. It contains 5 of the 44 unique subunits of complex I, and 5 of the 7 flavin domain proteins (NDUFV3, NDUFV2, NDUFV1, NDUFS1 and NDUFA2), with NDFS6 and NDUFA12 migrating on the gel at slice 55 and 41 respectively.

Of the known complex I assembly factors considered (NDUFAF1, NDUFAF2, NDUFAF3, NDUFAF4, NDUFAF5, NDUFAF6, FOXRED1, NUBPL, ACAD9, ECSIT, TMEM70, TMEM126A, TMEM126B and TIMMDC1), none were found to cluster with complex I in *Ndufs4* ($^{+/+}$) membranes (5.5a). For *Ndufs4* ($^{-/-}$) membranes, only NDUFAF2 and ACAD9 appear to co-migrate with the 830 kDa complex on the gel (5.5b). ACAD9 and NDUFAF2 are early and mid-late stage assembly factors, respectively (Andrews et al., 2013; Guerrero-Castillo et al., 2016): they are considered to dissociate from the complex I during attachment of the N-domain, consistent with their absence from both the full complex I and the flavin complex. A summary of peptide scores and coverages for the LC-MS analyses are given in (Table 5.2).

Ndufs4 ($^{-/-}$) complex I appears as three separate assemblies on BN-PAGE, an intact complex, an 830 kDa complex (without the flavin domain) and a flavin domain complex. In addition, two known complex I assembly factors, NDUFAF2 and ACAD9, were found to co-migrate with complex I. Therefore, in order to determine whether these observations are substantial or as a result of BN-PAGE artefact, complex I was purified and characterised from *Ndufs4* ($^{-/-}$) membranes.

a

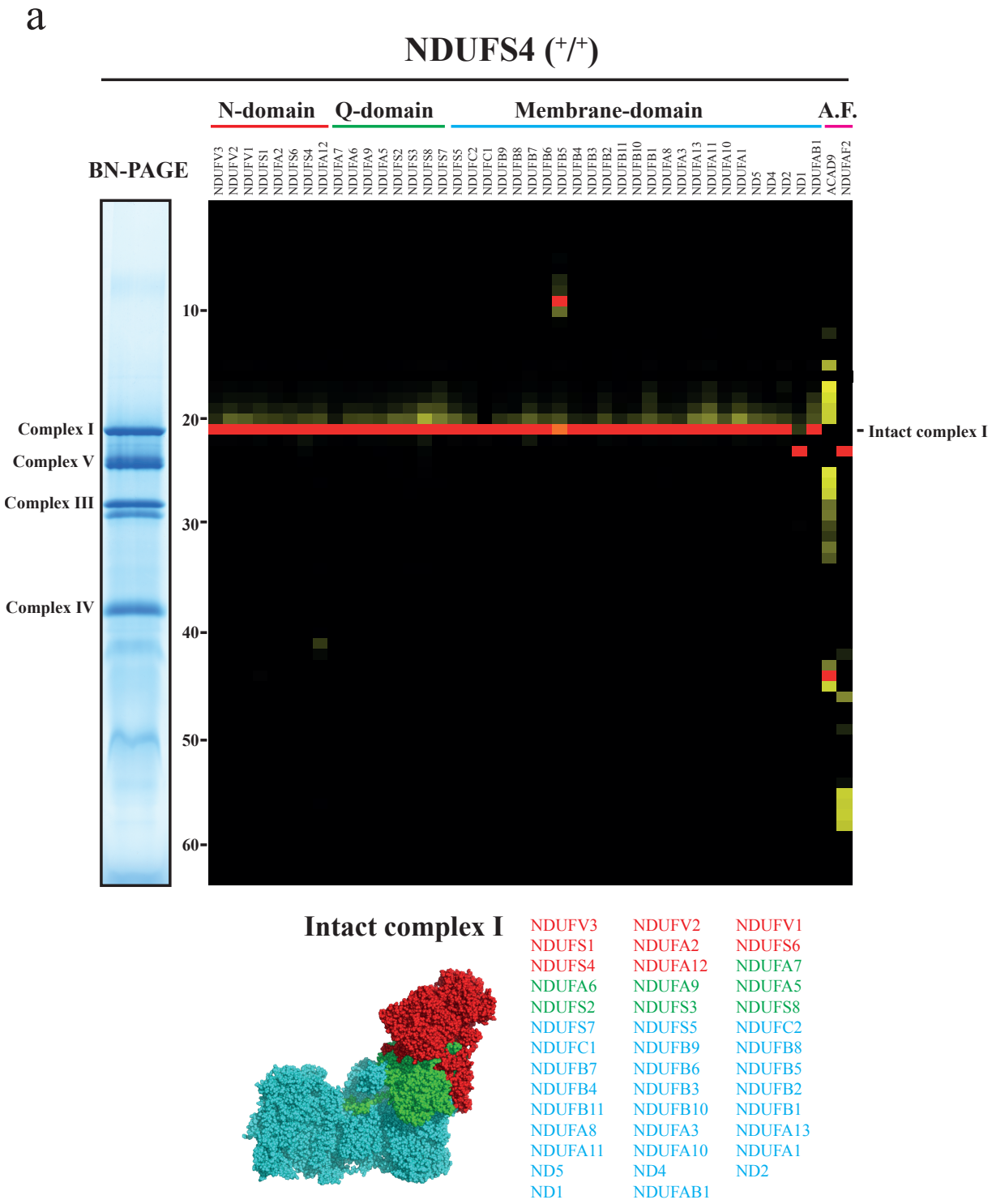


Figure 5.5: Complexome profiles of *Ndufs4* (-/-) and (+/+) solubilised membranes. 20 µg of solubilised *Ndufs4* (+/+) membranes (a) and *Ndufs4* (-/-) membranes (b) were separated by BN-PAGE and the gels stained with Coomassie Blue.

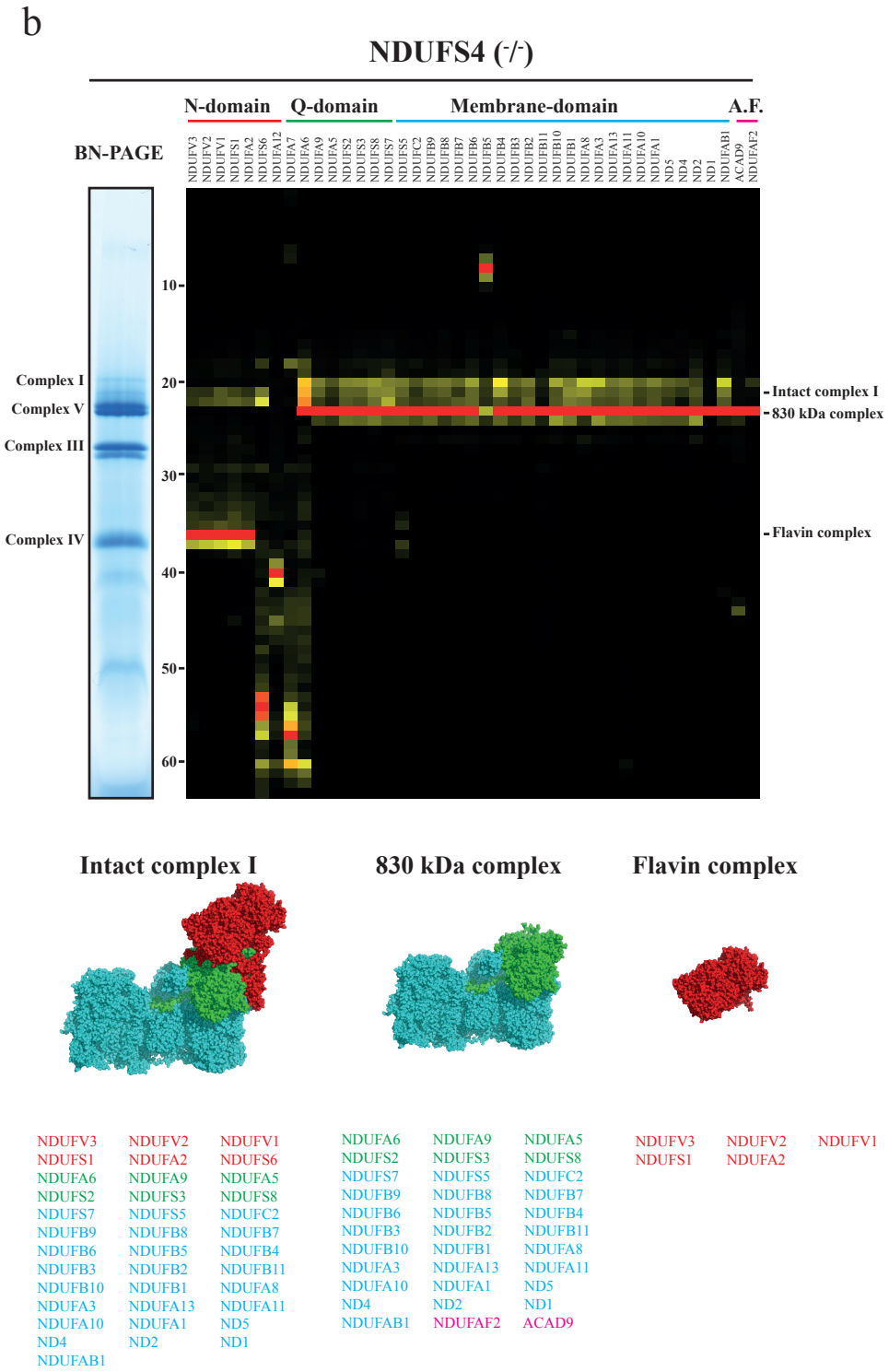


Figure 5.5: continued. The gel lanes were then cut into 64 equal-sized slices, followed by tryptic digestion and analyses by LC-MS. Hierarchal clustering of proteins was carried out according to previous work by Dr Mike Harbour (Heide et al., 2012). The heat map indicates the migration pattern of proteins and their relative abundance in the gel. Different domains of complex I detected in the analyses are coloured to correspond with the mouse active structure (Chapter 4). A.F. denotes assembly factors.

Subunit	*Ndufs4* ($^{+/+}$)			*Ndufs4* ($^{-/-}$)		
	Number of peptides	Sequence coverage	Score	Number of peptides	Sequence coverage	Score
NDUFA1	1	14%	136	1	14%	108
NDUFA2	7	55%	293	6	42%	274
NDUFA3	4	51%	241	4	51%	143
NDUFA5	6	44%	668	9	62%	772
NDUFA6	9	61%	635	8	56%	492
NDUFA7	7	55%	435	11	57%	513
NDUFA8	7	40%	395	8	40%	419
NDUFA9	21	74%	1420	22	77%	1530
NDUFA10	13	46%	972	16	53%	1845
NDUFA11	3	10%	256	2	9%	222
NDUFA12	10	57%	804	5	30%	260
NDUFA13	11	63%	1244	15	87%	817
NDUFAB1	5	46%	106	2	20%	115
NDUFB1	3	34%	122	2	32%	96
NDUFB2	2	12%	28	2	12%	46
NDUFB3	3	34%	244	4	34%	320
NDUFB4	8	54%	303	9	73%	308
NDUFB5	9	44%	462	8	44%	389
NDUFB6	11	65%	488	9	73%	370
NDUFB7	5	34%	947	6	45%	893
NDUFB8	6	51%	392	6	51%	586
NDUFB9	7	53%	447	7	45%	431
NDUFB10	8	45%	677	8	45%	549
NDUFB11	7	59%	838	7	48%	680
NDUFC1	1	78%	134	-	-	-
NDUFC2	8	54%	269	9	51%	269
NDUFS1	29	43%	2343	35	54%	2868
NDUFS2	19	61%	1573	19	64%	1766
NDUFS3	14	50%	814	13	48%	740
NDUFS4	8	65%	400	-	-	-
NDUFS5	4	39%	266	5	39%	277

NDUFS6	4	39%	479	5	46%	318
NDUFS7	9	36%	390	7	21%	512
NDUFS8	7	32%	278	9	46%	448
NDUFV1	14	33%	731	20	40%	796
NDUFV2	11	53%	1165	10	48%	912
NDUFV3	4	36%	182	4	38%	204
ND1	4	11%	367	3	11%	289
ND2	3	15%	279	3	14%	261
ND3	-	-	-	-	-	-
ND4	7	22%	663	5	13%	512
ND4L	-	-	-	-	-	-
ND5	6	12%	880	7	14%	422
ND6	-	-	-	-	-	-
ACAD9	7	13%	330	9	16%	428
NDUFAF2	2	13%	56	13	72%	425

Table 5.2: LC-MS peptide-based protein identifications for complexome profiling (Figure 5.5). The number of peptides detected from each subunit is given together with the sequence coverage of the mature protein. The score given for each subunit is the sum of the peptide score with denominator value of 25, representing the 95% confidence threshold ($p < 0.05$).

5.5 Purification of complex I from *Ndufs4* ($^{-/-}$) mouse membranes

Complex I from *Ndufs4* ($^{-/-}$) membranes was enriched using an anion exchange column according to the protocol described for *Ndufs4* ($^{+/+}$) complex I in Chapter 2.7 First, membranes were solubilised in 1% DDM, centrifuged, then the supernatant was loaded onto a HiTrap™ Q HP column. The complexes were eluted by increasing the salt concentration and complex I-containing fractions (Anion-exchange elution profile of *Ndufs4* ($^{-/-}$) complex I was initially determined by Western blotting: data not shown) were collected and concentrated for further purification (Figure 5.6). Interestingly, the *Ndufs4* ($^{-/-}$) complex I elutes in 350 mM NaCl, similar to *Ndufs4* ($^{+/+}$) complex I, so the subunit deletion has little effect on the anion-exchange elution profile of the enzyme.

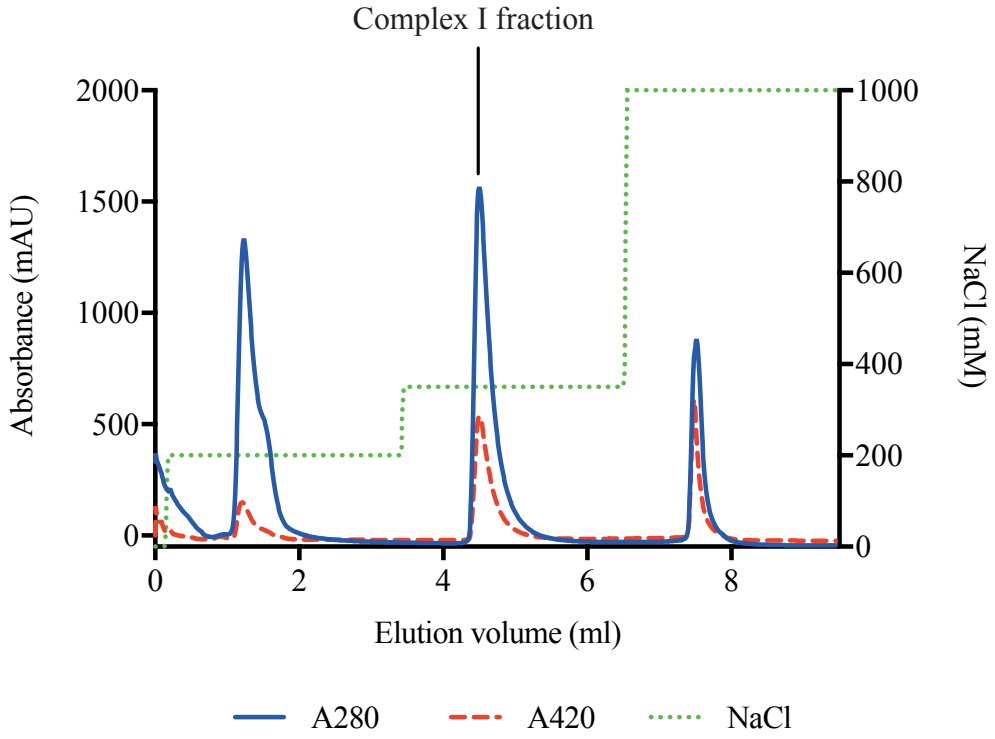


Figure 5.6: Purification of *Ndufs4* ($^{-/-}$) mouse complex I using anion-exchange chromatography. The Äkta micro system and the attached 1 mL HiTrap™ Q HP column were washed and equilibrated at a flow rate of 0.3 mL min^{-1} with buffer A (20 mM Tris-HCl pH 7.14 at 23 °C, 1 mM EDTA, 0.1% DDM, 10% ethylene glycol, 0.005% asolectin and 0.005% CHAPS), and proteins were eluted by increasing the proportion of buffer B (buffer A + 1 M NaCl).

In order to purify complex I further, size-exclusion chromatography was carried out, by injecting the pooled and concentrated complex I-containing fractions onto a Superose 6^{TM} Increase 3.2/300 column (Figure 5.7a). Complex I elutes at ~1.39 mL, with the tail of the peak overlapping with complex IV (main peak at ~1.64 mL). A significant portion of complex IV also appears on the shoulder of the main peak (~1.52 mL), presumably a dimer of the complex. The relative magnitude of the complex I peak compared to the complex IV peak is reduced two-fold for *Ndufs4* ($^{-/-}$) enzyme compared to *Ndufs4* ($^{+/+}$) enzyme. This is suggestive of a decrease in complex I content in the knockout, in agreement with BN-PAGE analyses (Figure 5.7). Moreover, there is a slight shift in the elution volume between *Ndufs4* ($^{+/+}$) complex I (1.36 mL) and *Ndufs4* ($^{-/-}$) complex I (1.39 mL), which may be cautiously attributed to a change in the hydrodynamic radius of the enzyme as a result of the subunit deletion. Otherwise, the purification of *Ndufs4* ($^{-/-}$) complex I is identical to that of *Ndufs4* ($^{+/+}$) complex I, indicating *Ndufs4* ($^{-/-}$) complex I is intact. This is in contradiction to the complexome analysis, with three complex I assemblies detected in *Ndufs4* ($^{-/-}$) complex I membranes (Figure 5.5b). To assess this further, purified *Ndufs4* ($^{-/-}$) complex I was analysed by BN-PAGE (5.7b), and it was found that the complex fragments under those conditions. The most prominent complex I band in the native gel corresponds in mass to the 830 kDa complex detected during the complexome analysis (see the asterisk in 5.7b).

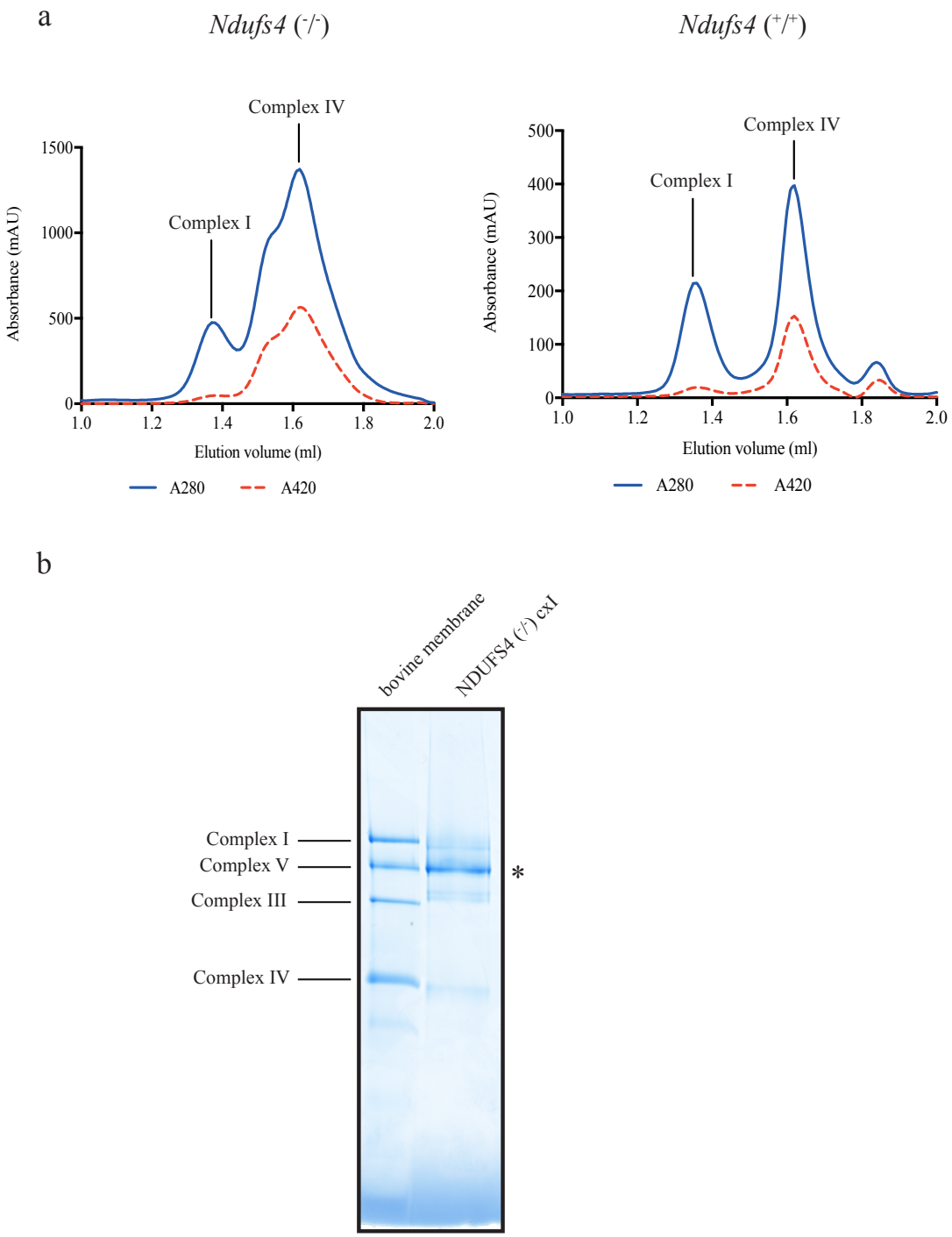


Figure 5.7: Purification of *Ndufs4* (-/-) mouse complex I using size-exclusion chromatography. (a) The Äkta micro system was operated at a flow rate of 0.03 mL min^{-1} with buffer C (20 mM Tris-HCl pH 7.14 at 23 °C, 10% ethylene glycol, 150 mM NaCl and 0.02% DDM), using a Superose 6^{TM} Increase 3.2/300 column. The purification trace for *Ndufs4* (-/-) complex I is compared to *Ndufs4* (+/+) complex I from Figure 3.3 (b *Ndufs4* (-/-) complex I (5 µg) and bovine membranes were analysed by BN-PAGE (see Chapter 2.5.1). The asterisk corresponds to the 830 kDa complex detected in the complexome analyses (5.5b)

5.6 Cross-linking *Ndufs4* ($^{-/-}$) mouse complex I in membranes

Due to the contradictory results obtained, where *Ndufs4* ($^{-/-}$) complex I behaves as an intact complex during size-exclusion chromatography (Figure 5.7) but not in BN-PAGE analyses (Figure 5.5b), it was proposed that complex I is disassembling during BN-PAGE. To investigate the proposal, a water-soluble cross-linker (BS^3) that reacts with primary amines (lysines and N-terminal amines) using both ends of its structure, releasing an N-hydroxysulphosuccinimide leaving group (Figure 5.8) (Shi et al., 2017). The BS^3 cross-linker was chosen because it is specific and the cross-linking distance (~11 Å) is ideal for intra-complex stabilisation, but not inter-complex cross-linking. Membranes were solubilised as usual with DDM detergent at 1%, then BS^3 was added and then the complexes analysed by BN-PAGE as described in Chapter 2.8 (Figure 5.9).

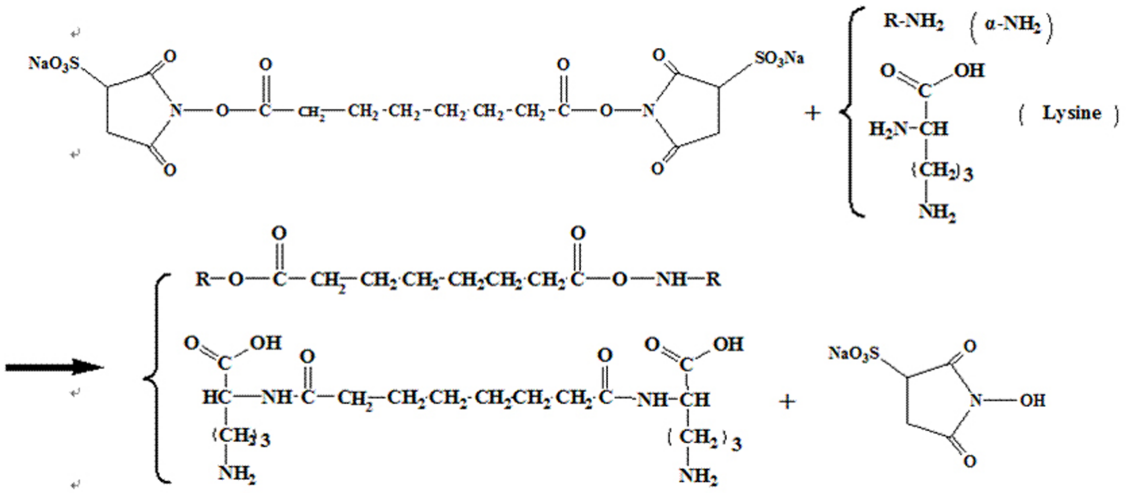


Figure 5.8: Mechanism of BS^3 cross-linker. Reaction scheme of BS^3 with Lysine. This figure was taken from Shi *et al.,* (Shi et al., 2017).

Alongside *Ndufs4* ($^{-/-}$) membranes, *Ndufs4* ($^{+/+}$) membranes were also cross-linked to confirm that the effects observed are not due to a native gel artefact (Figure 5.9). For *Ndufs4* ($^{-/-}$) membranes the amount of intact complex I detected by NADH-sensitive in-gel activity assay increases with higher concentrations of cross-linker used. Conversely, *Ndufs4* ($^{+/+}$) complex I remains the same with increasing concentration of cross-linker. However, there is an increase in the appearance of an NADH:NBT reactive band above complex I with increasing cross-

linker concentration (Figure 5.9). The identity of this protein is not known. Interestingly, with increasing cross-linker concentration, there follows the disappearance of a relatively low molecular weight band with in-gel activity (marked with an asterisk), concurrent with the increased detection of complex I. This lower molecular weight band corresponds to the flavin complex detected in the complexome analysis (Figure 5.5b). In opposition to previous results (Antonicka et al., 2003; Calvaruso et al., 2012; Ugalde et al., 2004; Valsecchi et al., 2012), the data presented here show that the 830 kDa complex is a result of destabilisation of *Ndufs4* ($^{-/-}$) complex I during BN-PAGE, and not a late-stage complex I assembly intermediate. The *Ndufs4* ($^{-/-}$) complex I is therefore intact but unstable.

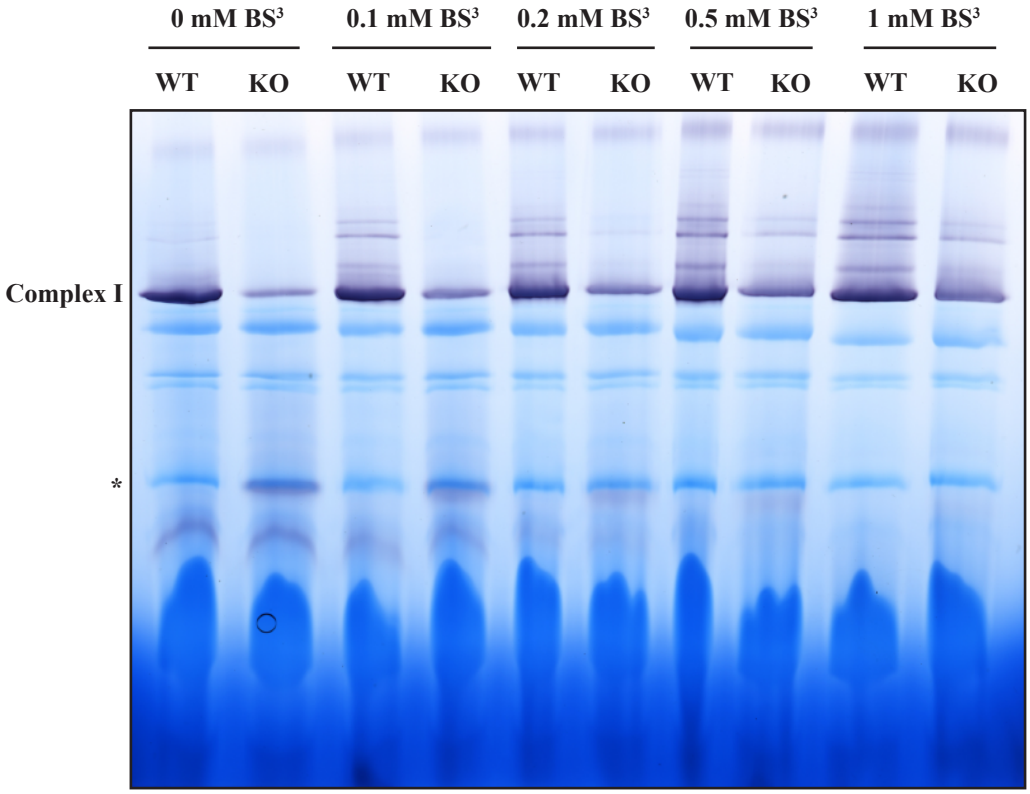


Figure 5.9: BN-PAGE in-gel complex I activity assay of cross-linked solubilised mouse membranes. Solubilised membranes (30 µg) incubated in various concentrations of BS^3 cross-linker were separated by BN-PAGE, and complex I was stained with 0.5 mg mL^{-1} and 100 µM NADH in 20 mM Tris-HCl pH 7.55 at 32 °C. *Ndufs4* ($^{+/+}$) and *Ndufs4* ($^{-/-}$) solubilised membranes are indicated by WT and KO, respectively. The asterisk indicates the position of the flavin domain of complex I.

5.7 Protein analyses of purified *Ndufs4* ($^{-/-}$) mouse complex I

To determine whether *Ndufs4* ($^{-/-}$) complex I has any assembly factors bound, or whether any subunits were dissociated during purification, as is the case for BN-PAGE, the protein content was analysed by mass spectrometry (the methods of which are described in Chapter 2.9).

Purified *Ndufs4* ($^{-/-}$) complex I was separated by SDS-PAGE, and then stained with Coomassie R-250 to visualise the proteins (Figure 5.10). The gel lane was cut into equal-size slices and then digested with trypsin. The resulting peptides were used to identify proteins by peptide mass fingerprinting using a MALDI TOF/TOF mass spectrometer (Figure 5.10 and Table 5.3). 26/43 unique subunits of *Ndufs4* ($^{-/-}$) complex I were identified using this method, with 8 and 18 being core and supernumerary subunits respectively. This is in contrast to *Ndufs4* ($^{+/+}$), where 33 subunits of complex I were identified (Figure 3.6 and Table 3.5). The subunits that are detected in *Ndufs4* ($^{+/+}$) complex I, but not in *Ndufs4* ($^{-/-}$) complex I, are NDUFAB1, NDUFB5, NDUFB6, NDUFB7, NDUFS5, NDUFS8, NDUFV2 and NDUFV3 (Table 5.3). These subunits are distributed in both the membrane and peripheral domains. Therefore, a lack of detection of those subunits probably reflects a general destabilisation of the enzyme. The assembly factors ACAD9 and NDUFAF2 were both identified, however, the sequence coverage of the former is low. Taking into account that ACAD9 is rich in suitable tryptic peptides, its low sequence coverage suggests it is a minor component in this preparation. Further protein contaminants in the preparation were also detected by MALDI TOF/TOF mass spectrometry, (Table 5.4), but are not considered relevant because they are not complex I-associated proteins.

To confirm the results obtained from the MALDI TOF/TOF analyses, the purified sample was subjected to tryptic digestion followed by analyses using LC-MS. 38 complex I subunits were detected, of which 10 were core subunits and 28 were supernumerary subunits (Table 5.3). The absence of particular subunits (NDUFAB1, NDUFB5, NDUFB6, NDUFB7, NDUFS5, NDUFS8, NDUFV2 and NDUFV3) in the MALDI TOF/TOF analyses but their detection by LC-MS suggests that they may be substoichiometric components, consistent with the *Ndufs4* ($^{-/-}$) complex I being destabilised. Assembly factors ACAD9 and NDUFAF2 were also detected, with low and high sequence coverage respectively. However, it is possible that the detected assembly factors are contaminants and not complex I-associated proteins.

Finally, in order to observe whether any complex I subunit modifications have changed as a result of the subunit deletion, subunit masses were measured by ESI-MS. Assignment of the masses to subunits of complex I were aided by previous bovine work (Carroll et al., 2009) and *Ndufs4* ($^{+/+}$) complex I measurements (Figure 3.7 and Table 3.7). 34 complex I subunits *Ndufs4* ($^{-/-}$) were assigned to their respective masses (Figure 5.11 and Table 5.5). There are no differences in the post-translation modifications between complex I from both strains. Masses corresponding to ACAD9 and NDUFAF2 were not measured, suggestive that these are either minor components in the preparation or they are more difficult to detect.

Using mass spectrometry, 41/43 unique subunits of *Ndufs4* ($^{-/-}$) complex I were identified, consistent with complex I being intact. However, the lower detection of several subunits suggests the enzyme is unstable and loses some of its subunits during purification. Although NDUFAF2 was detected in the purified complex I in Δ*Ndufs4 Y. lipolytica* (Kahlhöfer et al., 2017), mass spectrometry evidence for its association with mouse NDUFS4 ($^{-/-}$) complex I is not convincing.

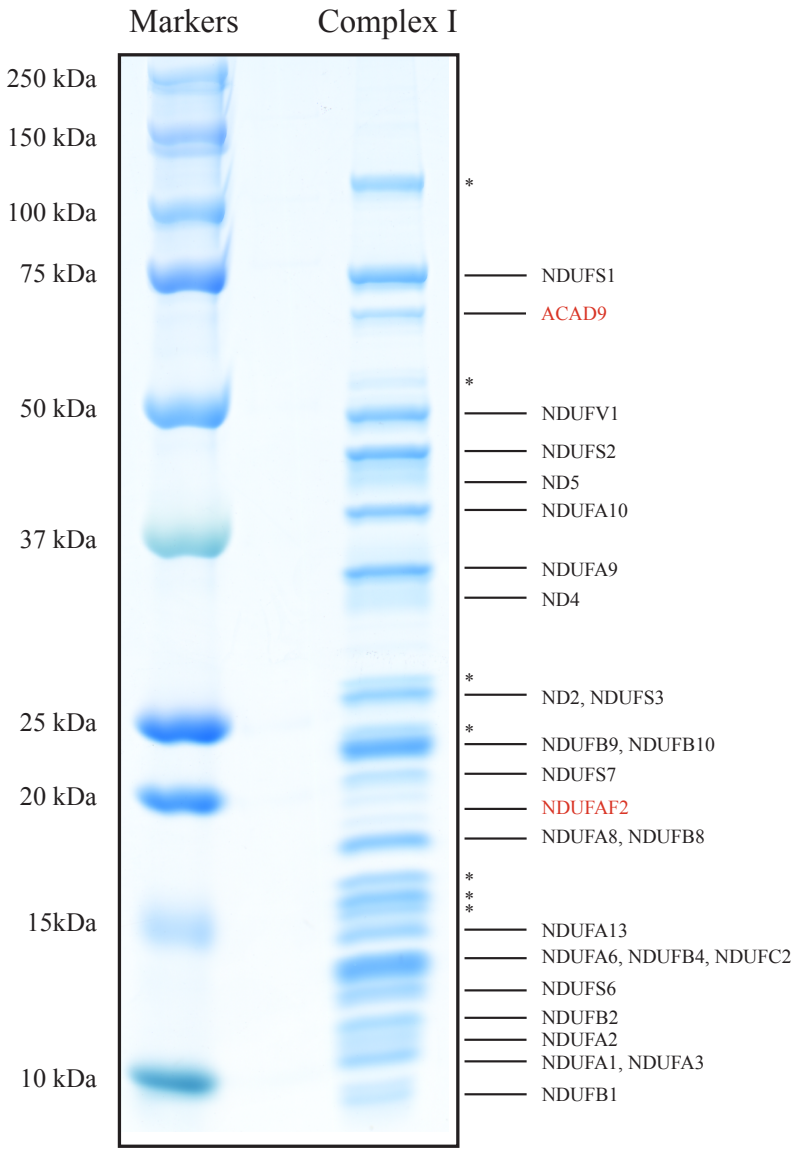


Figure 5.10: SDS-PAGE analysis of purified *Ndufs4* ($^{-}/^{-}$) mouse complex I. Complex I subunits (5 µg) were separated by SDS-PAGE and then stained by Coomassie Blue. Subunits were identified by peptide mass fingerprinting using a MALDI-TOF/TOF mass spectrometer. Assembly factors are displayed in red text. * Indicates proteins not considered to be associated with complex I (see Table 5.4). Known molecular size markers were loaded for reference.

Subunit	Number of residues	MALDI TOF/TOF MS/MS			LC MS		
		Number of peptides	Sequence coverage	Score	Number of peptides	Sequence coverage	Score
NDUFA1	70	2	18%	52/33	6	42%	140
NDUFA2	98	5	38%	243/32	9	71%	212
NDUFA3	83	3	32%	40/33	4	51%	97
NDUFA5	115	3	31%	109/33	10	64%	433
NDUFA6	130	3	26%	148	19	88%	710
NDUFA7	112	-	-	-	18	86%	756
NDUFA8	171	4	26%	127/32	12	52%	423
NDUFA9	342	11	39%	596/34	25	71%	1331
NDUFA10	320	11	46%	600/33	20	45%	896
NDUFA11	140	-	-	-	2	10%	194
NDUFA12	145	-	-	-	14	83%	615
NDUFA13	143	5	30%	230/33	20	82%	1426
NDUFAB1	88	-	-	-	1	20%	71
NDUFB1	56	4	62%	123/33	5	51%	82
NDUFB2	72	2	12%	70/32	1	12%	41
NDUFB3	103	2	14%	39/33	5	33%	115
NDUFB4	128	7	53%	302/32	17	70%	601
NDUFB5	143	-	-	-	9	46%	412
NDUFB6	127	-	-	-	15	76%	404
NDUFB7	136	-	-	-	5	45%	479
NDUFB8	157	5	45%	273/32	9	51%	550
NDUFB9	178	3	24%	74/32	16	66%	467
NDUFB10	175	1	8%	54/32	11	55%	520
NDUFB11	122	-	-	-	11	69%	521
NDUFC1	49	-	-	-	-	-	-
NDUFC2	120	3	17%	126/32	12	60%	359
NDUFS1	704	9	13%	227/33	44	48%	2754
NDUFS2	430	11	33%	479/34	23	52%	1063
NDUFS3	228	6	30%	336/33	18	53%	978
NDUFS4	133	-	-	-	-	-	-
NDUFS5	105	-	-	-	10	46%	244

NDUFS6	96	4	39%	189/33	6	60%	518
NDUFS7	189	5	17%	199/33	10	36%	441
NDUFS8	178	-	-	-	9	40%	228
NDUFV1	444	10	27%	475/34	24	43%	1039
NDUFV2	217	-	-	-	13	51%	955
NDUFV3	69	-	-	-	6	68%	142
ND1	318	-	-	-	3	8%	140
ND2	345	1	4%	61/32	-	-	-
ND3	115	-	-	-	-	-	-
ND4	459	3	7%	112/34	3	7%	118
ND4L	98	-	-	-	-	-	-
ND5	607	1	2%	50/32	5	8%	253
ND6	172	-	-	-	-	-	-
ACAD9	625	2	4%	54/33	4	8%	196
NDUFAF2	168	4	23%	157/34	11	71%	342

Table 5.3: Peptide-based protein identification of the composition of *Ndufs4* ($^{-}/^{-}$) complex I.

The number of peptides detected from each subunit is given together with the sequence coverage of the mature protein (number of residues column). For the two analytical methods, the score given for each of the subunits is the sum of the peptide score with denominator (denominator value for LC-MS score is 24) representing the 95% confidence threshold ($p < 0.05$).

Protein	Number of residues	MALDI TOF/TOF MS/MS		
		Number of peptides	Sequence coverage	Score
OGDH	1019	7	9%	265/33
SLC4A1	929	4	5%	193/33
HSPA9	679	3	6%	98/32
ACADVL	634	10	22%	407/33
ATP5A1	503	8	24%	379/33
DLD	509	3	4%	168/32
ATP5B	529	6	17%	221/32
DLST	454	3	4%	62/32
COX1	514	1	1%	49/33
VDAC1	283	3	20%	83/34
PHB	272	6	28%	135/32
COX4I1	169	6	30%	208/34
GM11273	129	2	14%	88/33
COX6C	76	4	36%	135/34
COX6B1	86	2	20%	86/32
COX7A1	80	2	28%	113/33
COX7C	63	2	28%	96/33

Table 5.4: Identification of contaminants in *Ndufs4* ($^{-}$/$^{-}$) preparation. The number of peptides detected from each subunit is given together with the sequence coverage of the protein. The score given for each of the subunits is the sum of the peptide score with denominator representing the 95% confidence threshold ($p < 0.05$).

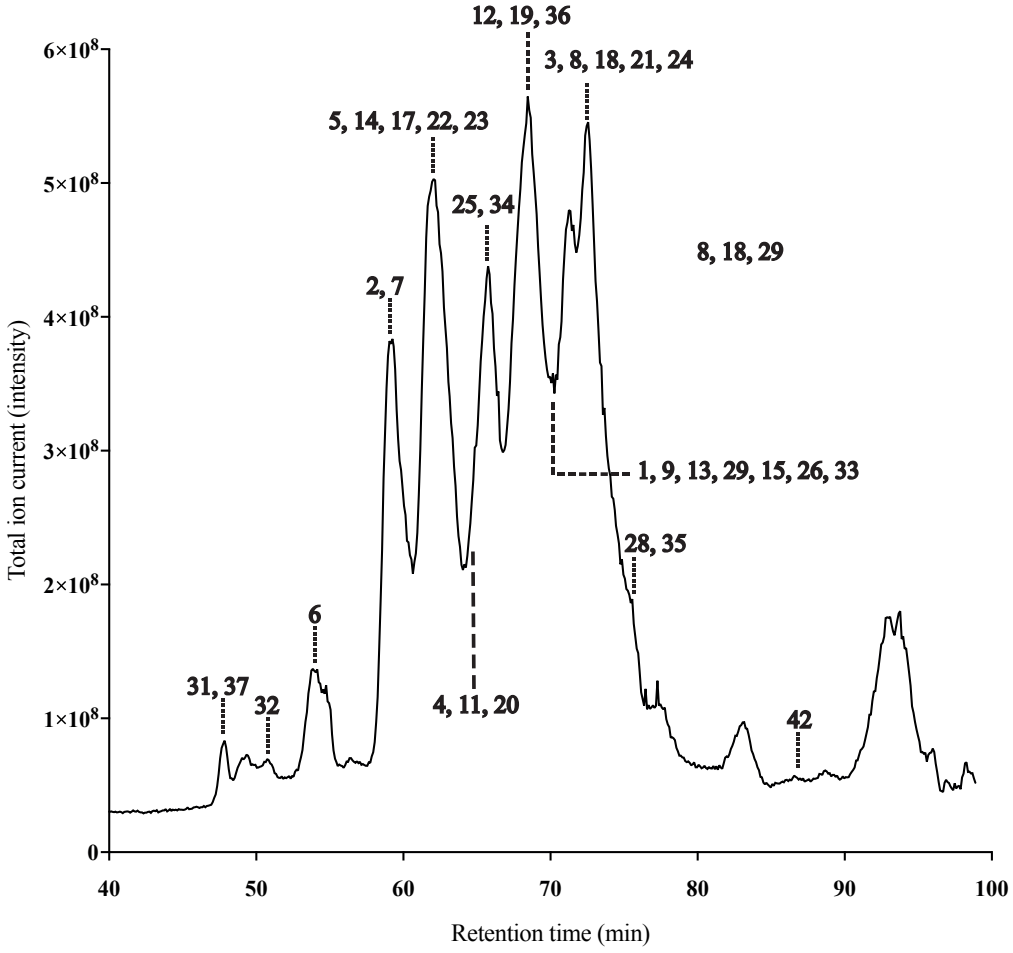


Figure 5.11: Analysis of *Ndufs4* ($^{-}/^{-}$) mouse mitochondrial complex I by LC-MS. Complex I (30 µg**)** subunits were separated by reverse-phase HPLC and detected by ESI-MS. The ion intensities of the eluents are indicated, and complex I subunits are identified and numbered according to Table 5.5.

Subunit	Mass		Δ mass discrepancy	Modifications
	Calculated	Observed		
1. NDUFA1	8138.5	8139.6	1.1	-
2. NDUFA2	10826.5	10826.2	-0.3	-Met, +N-acetyl
3. NDUFA3	9241.7	9242.5	0.8	-Met, +N-acetyl
4. NDUFA5	13270.5	13270.7	0.2	-Met, +N-acetyl
5. NDUFA6	15193.6	15193.8	0.2	-Met, +N-acetyl
6. NDUFA7	12486.4	12486.4	0.0	-Met, +N-acetyl
7. NDUFA8	19852.9	19854.8	1.9	-Met, +4×Cys-Cys
8. NDUFA9	38825.8	38841.0	15.2	-ΔImport (1-35)
9. NDUFA10	36943.0	36948.0	5.0	-ΔImport (1-35)
10. NDUFA11	14851.1	-	-	-
11. NDUFA12	17128.4	17127.7	-0.7	+N-acetyl
12. NDUFA13	16770.4	16770.8	0.4	-Met, +N-acetyl
13. NDUFAB1	10662.1	10661.9	-0.2	-ΔImport (1-68), +4PP-acyl chain
14. NDUFB1	6822.9	6823.3	0.4	-Met
15. NDUFB2	8491.4	8491.2	-0.2	-ΔImport (1-33)
16. NDUFB3	11561.0	-	-	-
17. NDUFB4	14992.1	14992.0	-0.1	-Met, +N-acetyl
18. NDUFB5	16858.4	16858.6	0.2	-ΔImport (1-46)
19. NDUFB6	15425.9	15426.5	0.6	-Met, +N-acetyl
20. NDUFB7	16405.8	16407.2	1.4	-Met, +N-myristoyl, +2×Cys-Cys
21. NDUFB8	18745.1	18746.4	1.3	-ΔImport (1-29)
22. NDUFB9	21894.9	21895.7	0.8	-Met, +N-acetyl
23. NDUFB10	20888.6	20890.7	2.1	-Met, +2×Cys-Cys
24. NDUFB11	14343.0	14343.6	0.6	-ΔImport (1-29)
25. NDUFC1	5752.6	5753.8	1.2	-ΔImport (1-27)
26. NDUFC2	14205.7	14207.2	1.5	+N-acetyl
27. NDUFS1	77210.6	-	-	-
28. NDUFS2	49257.6	49276.4	18.8	ΔImport (1-27), +dimethyl-R85
29. NDUFS3	26479.0	26479.6	0.6	-ΔImport (1-35)
30. NDUFS4	15325.3	-	-	-ΔImport (1-42)

31. NDUFS5	12512.5	12513.3	0.8	-Met, +2×Cys-Cys
32. NDUFS6	10777.0	10775.6	-1.4	-ΔImport (1-20)
33. NDUFS7	21032.5	21036.3	3.8	-ΔImport (1-35), +Hydroxyl-R87
34. NDUFS8	20442.2	20438.6	-3.6	-ΔImport (1-35)
35. NDUFV1	48626.5	48646.3	19.8	-ΔImport (1-20)
36. NDUFV2	23847.4	23846.5	-0.9	-ΔImport (1-31)
37. NDUFV3	8027.8	8028.5	0.7	-ΔImport (1-35)
38. ND1	36087.3	-	-	-
39. ND2	38780.4	-	-	-
40. ND3	13246.9	-	-	-
41. ND4	51909.8	-	-	-
42. ND4L	10634.7	10634.3	-0.4	+N-formyl
43. ND5	68502.8	-	-	-
44. ND6	18654.2	-	-	-

Table 5.5: Molecular mass measurement of complex I subunits. Measured masses were assigned to complex I subunits according to the mass discrepancy between the observed masses and calculated masses (mature protein with known modifications), as shown in previous work (Carroll et al., 2009) and in Chapter 3. Modification of the observed masses include (-) cleavage of import sequences or residues, or (+) or addition of post-translation modifications. Disulphide bridges are presented as Cys-Cys.

5.8 Association of FMN to *Ndufs4* ($^{-/-}$) mouse complex I

It was previously established in this thesis that *Ndufs4* ($^{-/-}$) complex I was intact but unstable, and a loss of subunits may occur during purification, therefore the FMN content of the enzyme was investigated. Flavin contents of both *Ndufs4* ($^{+/+}$) complex I and *Ndufs4* ($^{-/-}$) complex I were measured, using known FMN standards as a reference (described in chapter 2.10). Purified *Ndufs4* ($^{-/-}$) complex I contains noticeably less FMN than its wild type counterpart enzyme (Table 5.6). However, the amount of FMN in *Ndufs4* ($^{+/+}$) complex I is not stoichiometric, due to the lack of sensitivity in this method, therefore it is not clear whether the lower amount of detected FMN in *Ndufs4* ($^{-/-}$) complex I is relevant.

FMN concentration (µM) in complex I	
Ndufs4 ($^{-/-}$)	*Ndufs4* ($^{+/+}$)
0.93 ± 0.02	1.30 ± 0.02

Table 5.6: Flavin content of mouse complex I. Details of the methods employed are described in chapter 2. The flavin concentration in complex I (1 µM) was measured by reference to a set of prepared standards ranging from 0.2 – 10 µg mL^{-1}. The data are presented as mean values with error bars for the standard error of the mean (n ≥ 3).

The association of FMN to complex I was then probed. Complex I was exposed to gradually increasing temperature, which leads to the dissociation of FMN (Figure 5.12). The temperature at which 50% of the FMN has dissociated can then be calculated (Tm_{50}). The Tm_{50} for the purified knockout enzyme is 52.3 °C, which is essentially the same as that of the purified enzyme (51.2 °C, Figure 3.4). This suggests that there is no difference between the strength of association of FMN with complex I between *Ndufs4* ($^{-/-}$) complex I and *Ndufs4* ($^{+/+}$). Therefore, instability of *Ndufs4* ($^{-/-}$) complex I is the issue and not loss of FMN.

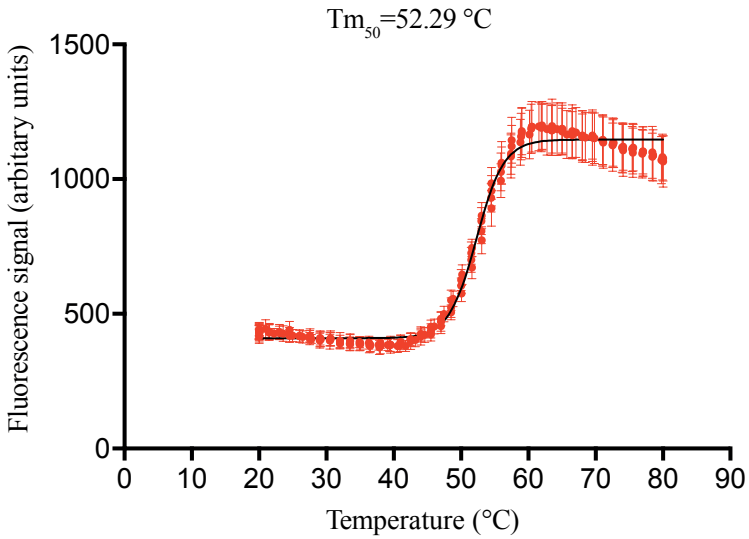
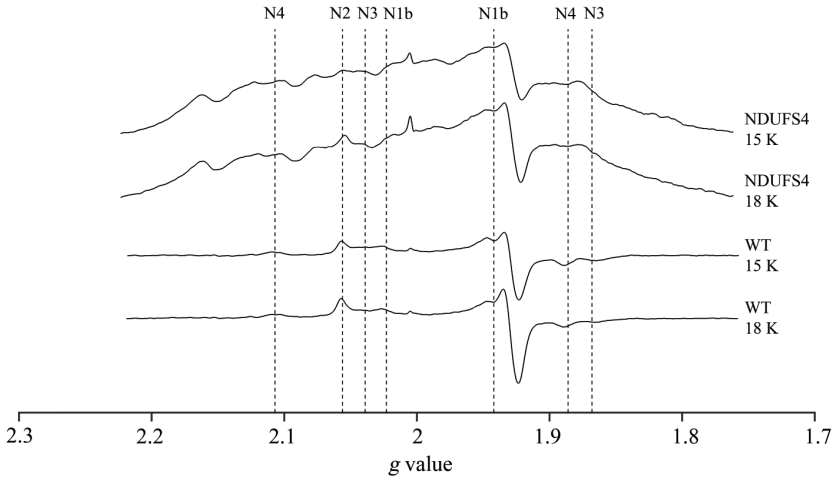


Figure 5.12: Thermal-based assay for flavin association with purified *Ndufs4* ($^{-/-}$) mouse complex I. Details of this assay are described in chapter 2. Purified complex I (6 µg) was exposed to increasing temperature from 20 – 80 °C, and the fluorescence signal of FMN was measured. The temperature at which complex I loses 50% FMN (Tm_{50}) was calculated (black sigmoidal curve).

5.9 EPR measurements on the iron-sulphur clusters in *Ndufs4* ($^{-/-}$) mouse complex I

To characterise the iron-sulphur clusters and establish whether the subunit deletion in complex I has affected its ability to transfer electrons, X-band continuous wave EPR spectroscopy was carried out on *Ndufs4* ($^{-/-}$) complex I and compared to *Ndufs4* ($^{+/+}$) complex I (Figure 5.13a). Purified *Ndufs4* ($^{-/-}$) complex I was reduced with NADH and then plunge frozen. Assignment of signals in the EPR spectra to particular iron-sulphur clusters were aided by comparison to the previous bovine enzyme (Reda et al., 2008) and the mouse enzyme (Figure 3.5a). The positions of assigned iron-sulphur clusters are shown in the structure for reference (Figure 5.13b). Although the signal to noise for the spectra was poor, 4 out of 8 iron-sulphur clusters were detected for *Ndufs4* ($^{-/-}$) complex I (N1b, N2, N3 and N4), consistent with the EPR spectra of *Ndufs4* ($^{+/+}$) complex I (Figure 5.13a). Cluster N5, which was not detected, is fast-relaxing when reduced, therefore it is difficult to obtain a signal from it (Reda et al., 2008a). However, the overall signal amplitude of the spectra is significantly diminished in *Ndufs4* ($^{-/-}$) complex I compared to *Ndufs4* ($^{+/+}$) complex I. This is because of difficulty encountered in obtaining sufficient *Ndufs4* ($^{-/-}$) complex I to give good EPR spectra. Interestingly, EPR spectra on the clusters of purified complex I from Δ*Ndufs4 Y. lipolytica* show broadening of the signals (Kahlhöfer et al., 2017), in particular N1b, which is in close proximity to *Ndufs4* in the wild type strain, suggesting a perturbation of the clusters as a result of subunit loss in complex I. In the case of mouse *Ndufs4* ($^{-/-}$) complex I, perturbation of the clusters is not clear because of poor signal to noise in the EPR spectra. Despite this, it can be inferred that *Ndufs4* ($^{-/-}$) complex I has all of its iron-sulphur clusters, although it is unclear whether their midpoint potentials have changed relative to *Ndufs4* ($^{+/+}$) complex I, thus affecting the rate of intra-molecular electron transfer.

a



b

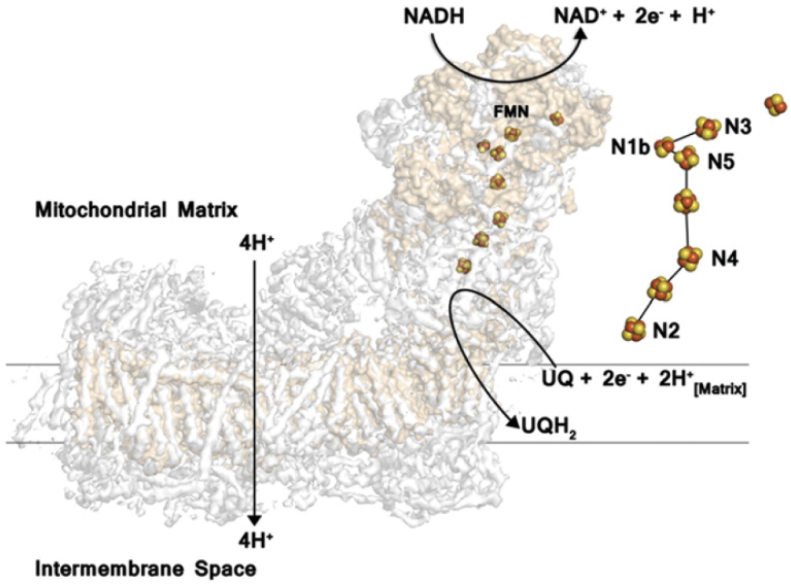


Figure 5.13: X-band continuous wave EPR spectra of *Ndufs4* ($^{-}/^{-}$) mouse complex I. (a) Details of the conditions of the EPR measurements are described in Chapter 2.11. 9 µl of purified *Ndufs4* ($^{-}/^{-}$) complex I (100 µg) and *Ndufs4* ($^{+}/^{+}$) complex I (180 µg) were reduced by 2 mM NADH and the first derivative of the absorbance of reduced iron-sulphur clusters (labelled) were detected by varying the temperature (15 K and 18 K) and magnetic field, but the microwave frequency was constant (~9 GHz). The *Ndufs4* ($^{-}/^{-}$) complex I spectra was scaled by ~2× to adjust for the concentration difference between the complexes (b) The structure of the bovine complex I (brown: core subunits and grey: supernumerary subunits) is depicted with the iron-sulphur clusters (yellow-orange spheres). This figure taken from Wright et al (Wright et al., 2016).

5.10 Activity measurements on purified *Ndufs4* ($^{-/-}$) mouse complex I

It has been established that loss of NDUFS4 from complex I leads to multiple issues, such as low catalytic activity and instability of the enzyme. To further confirm this, activity measurements were carried out on isolated *Ndufs4* ($^{-/-}$) complex I, and compared to the previously characterised *Ndufs4* ($^{+/+}$) complex I (Table 5.7). The piericidin-sensitive NADH:DQ oxidoreduction rate of *Ndufs4* ($^{-/-}$) complex I (representative of the enzyme's catalytic turnover) was 7% the value of *Ndufs4* ($^{+/+}$) complex I. This is consistent with the results from other groups for *Ndufs4* ($^{-/-}$) complex I in mitochondria (Kruse et al., 2008; Petruzzella et al., 2001; Quintana et al., 2010), confirming its catalytic dysfunction.

To probe the source of the catalytic dysfunction, the flavin based oxidation of NADH was probed using artificial NADH oxidising substrates (HAR, APAD and FeCN). All the flavin activities were lower in *Ndufs4* ($^{-/-}$) complex I compared to *Ndufs4* ($^{+/+}$) complex I (Table 5.7). However, as previously observed in membranes, the relative activities of *Ndufs4* ($^{-/-}$) complex I compared to *Ndufs4* ($^{+/+}$) complex I differed substantially, with 7% for APAD, 16% for FeCN and 44% for HAR. Further biochemical characterisation of the flavin-site is required to understand the underlying reason for these rate differences in NADH oxidation, but it may be due to instability, change in the mechanism of NADH oxidation or possibly perturbation in the iron-sulphur clusters.

Assay	*Ndufs4* ($^{-/-}$) (μmol min^{-1} mg^{-1})	*Ndufs4* ($^{+/+}$) (μmol min^{-1} mg^{-1})	Relative activities (%)
NADH:DQ	0.81 ± 0.03	11.99 ± 0.11	7
H_2O_2	8.77 ×10^{-3} ± 0.26 ×10^{-3}	15.30 ×10^{-3} ± 0.19 ×10^{-3}	57
FeCN	5.10 ± 0.06	32.87 ± 1.74	16
HAR	21.06 ± 0.16	47.38 ± 1.49	44
APAD	0.29 ± 0.01	4.46 ± 0.09	7

Table 5.7 Kinetic characterisation of purified *Ndufs4* ($^{-/-}$) and *Ndufs4* ($^{+/+}$) mouse complex I. Rates for the purified enzyme were measured according to the protocols in chapter 2. Values for rates of *Ndufs4* ($^{+/+}$) were taken from Table 3.4. The data are presented as mean values with error bars for the standard error of the mean (n ≥ 3).

It has been previously reported that NDUFS4 deletion is linked with an increase in the level of reactive oxygen species produced (Valsecchi et al., 2013), therefore H_2O_2 production was measured from both strains. However, H_2O_2 production by *Ndufs4* ($^{-/-}$) complex I is 57% of that of *Ndufs4* ($^{+/+}$) complex I. It is not clear whether the results are pathophysiologically relevant or is a consequence of *Ndufs4* ($^{-/-}$) complex I destabilisation during purification.

5.11 HAR titration on *Ndufs4* ($^{-/-}$) mouse membranes

To further probe the unusual kinetics of the flavin-mediated reactions of *Ndufs4* ($^{-/-}$) complex I, HAR was chosen for further investigation because of its comparatively high activity (Table 5.7). The rates of NADH oxidation by membranes from both strains were measured with various concentrations of HAR in the assay (Figure 5.14). The maximal rates (V_{max}) of NADH oxidation were comparable between *Ndufs4* ($^{-/-}$) and *Ndufs4* ($^{+/+}$) complex I (Figure 5.14a, b, Table 5.8). However, considering that the complex I content is less in NDUFS4 ($^{-/-}$) compared to *Ndufs4* ($^{+/+}$) membranes (diminished by at least half as observed by BN-PAGE in Figure 5.4), the rate of NADH oxidation per complex I in membranes is much greater for the *Ndufs4* ($^{-/-}$). Further analyses showed a striking 120-fold decrease in the K_M value for *Ndufs4* ($^{-/-}$) compared to *Ndufs4* ($^{+/+}$) membranes. K_M, a value describing substrate concentration at half-maximal turnover, is dependent on the maximal velocity of the enzyme and its binding affinity with the substrate(s). Since the V_{max} is increased in *Ndufs4* ($^{-/-}$) membranes compared to *Ndufs4* ($^{+/+}$) membranes, the decrease in K_M of *Ndufs4* ($^{-/-}$) membranes relative to *Ndufs4* ($^{+/+}$) membranes is attributable to an increased binding affinity of the substrate HAR to complex I. As HAR is thought to bind to the nucleotide phosphate in NADH (Birrell et al., 2011), it may be posited that binding of NADH may change as a result of structural changes in *Ndufs4* ($^{-/-}$) complex I, and consequently the binding of HAR to the enzyme, but to confirm this requires further biochemical characterisation of the flavin-site or the structure of the enzyme.

a b

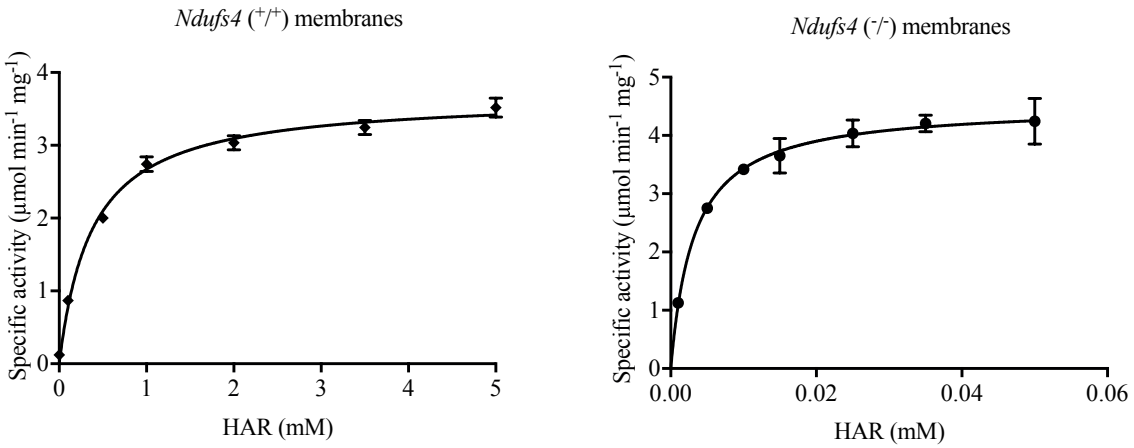


Figure 5.14: NADH:HAR activity measurements. Details of the protocol are described in Chapter 2. 4 µg mL^{-1} of membranes (a,b) were incubated with various concentrations of HAR. The reactions were initiated with 100 µM NADH. The data are presented as mean values with error bars for the standard error of the mean (n ≥ 3).

Samples	*Ndufs4* (-/-)		*Ndufs4* (+/+)	
	V_{max} (µmol min^{-1} mg^{-1})	K_m (mM)	V_{max} (µmol min^{-1} mg^{-1})	K_m (mM)
Membranes	4.52 ± 0.09	$3.2 \times 10^{-3} \pm 0.32 \times 10^{-4}$	3.67 ± 0.06	0.38 ± 0.03

Table 5.8: Michaelis–Menten parameters for NADH:HAR activity measurements. Values are from the experiments of figure 5.14, with the experimental protocols discussed in chapter 2. The data are presented as mean values with error bars for the standard error of the mean (n ≥ 3).

5.12 Purification of cross-linked *Ndufs4* (-/-) mouse complex I

In order to stabilise *Ndufs4* (-/-) complex I to prevent the loss of its components, solubilised membranes were cross-linked with 1 mM BS^3 as described in chapter 2.8 Otherwise, the purification procedure was identical to that used previously.

First, complex I-containing fractions were enriched by anion-exchange chromatography (Figure 5.15). Similar to the previous complex I purifications, cross-linked *Ndufs4* ($^{-/-}$) complex I eluted in 350 mM NaCl. There is a large peak at 200 mM NaCl, which is not detected in the non-cross-linked preparation (Figure 5.15), which is assumed to arise from unreacted BS^3. Complex I-containing fractions were pooled and concentrated for further separation as usual.

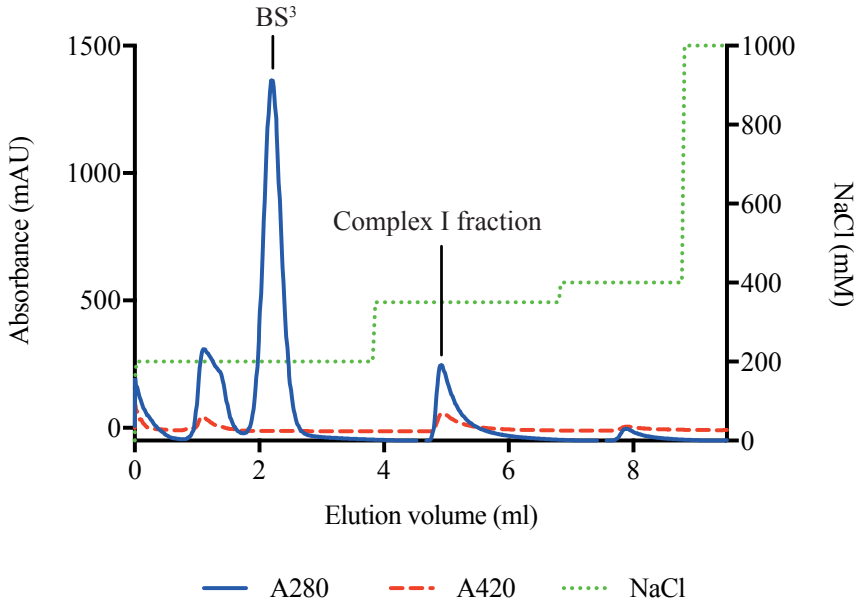


Figure 5.15: Purification of cross-linked *Ndufs4* ($^{-/-}$) mouse complex I using anion-exchange chromatography. The Äkta micro system and the attached 1 mL HiTrap™ Q HP column were washed and equilibrated at a flow rate of 0.3 mL min^{-1} with buffer A (20 mM Tris-HCl pH 7.14 at 23 °C, 1 mM EDTA, 0.1% DDM, 10% ethylene glycol, 0.005% asolectin and 0.005% CHAPS), and proteins were eluted in buffer B (buffer A + 1 M NaCl). Elution of BS^3 and complex I-containing fractions are shown.

To further resolve complex I, size-exclusion chromatography was used to separate complexes IV and I (Figure 5.16). In the size-exclusion trace, both complexes I and IV elution volume have shifted to higher elution volumes for the *Ndufs4* ($^{-/-}$) preparation compared to the non-cross-linked *Ndufs4* ($^{+/+}$) preparation (Figure 5.16), appearing smaller. It is plausible that cross-linking has affected the migration of both enzymes in the size-exclusion column, as evident by the equal shift in their elution. Complex I from this preparation displayed an NADH:DQ activity of 0.86 µmol min^{-1} mg^{-1}, meaning the enzyme is essentially catalytically inactive.

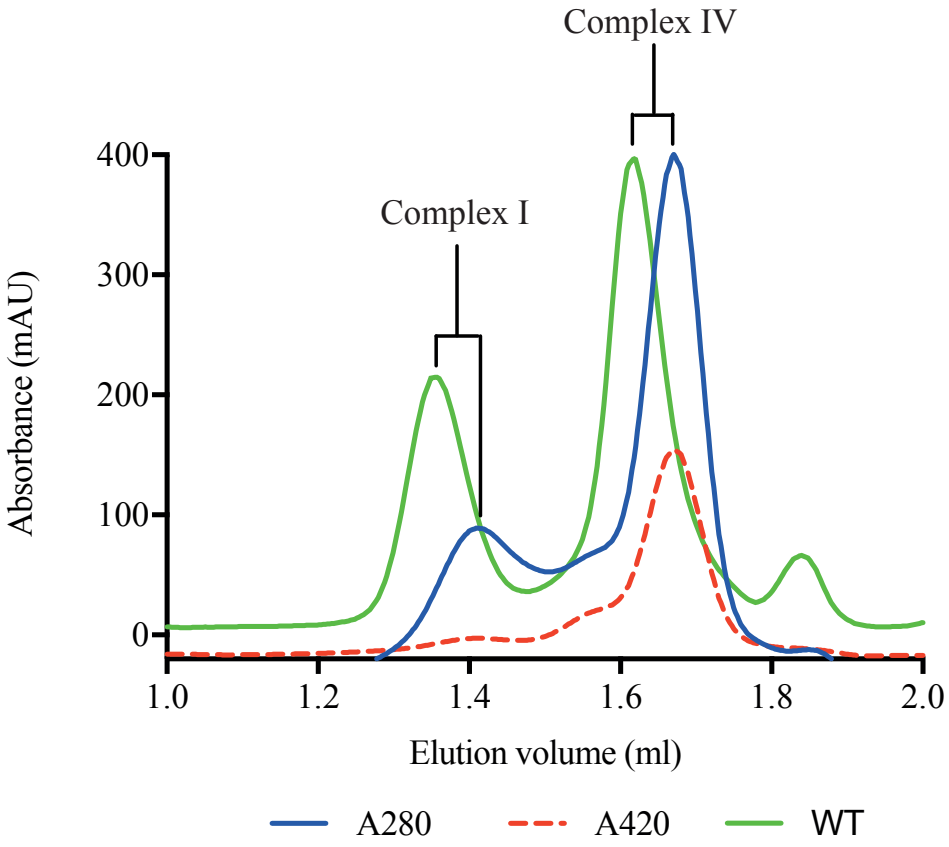


Figure 5.16: Purification of cross-linked *Ndufs4* ($^{-/-}$) mouse complex I using size-exclusion chromatography. The Äkta micro system was operated at a flow rate of 0.03 mL min^{-1} with buffer C (20 mM Tris-HCl pH 7.14 at 23 °C, 10% ethylene glycol, 150 mM NaCl and 0.05% DDM), using a Superose 6^{TM} Increase 3.2/300 column. The purification trace (280 nm) for non-crosslinked *Ndufs4* ($^{+/+}$) complex I (WT) from Figure 3.3 is shown as a reference.

The size-exclusion purified complex I was then assessed by BN-PAGE, to determine whether it was intact (Figure 5.17). The in-gel activity assay shows that the purified cross-linked *Ndufs4* ($^{-/-}$) complex I is indeed intact, corresponding in size to complex I in *Ndufs4* ($^{+/+}$) membranes. There is a slight shift on the gel in the migration of the *Ndufs4* ($^{-/-}$) complex I, which may be due to the cross-linking itself, as opposed to the loss of subunits. However, there is an NADH sensitive band migrating below the complex IV band in *Ndufs4* ($^{+/+}$) membranes. There is also minor band present where complex V migrates in *Ndufs4* ($^{+/+}$) membranes. These bands were not identified because cross-linking affects tryptic digestion, thus complicating identification of protein components in the preparation. Nevertheless, because the majority of purified enzyme is NADH sensitive and corresponds to the size of complex I, it can be stated that it is largely intact.

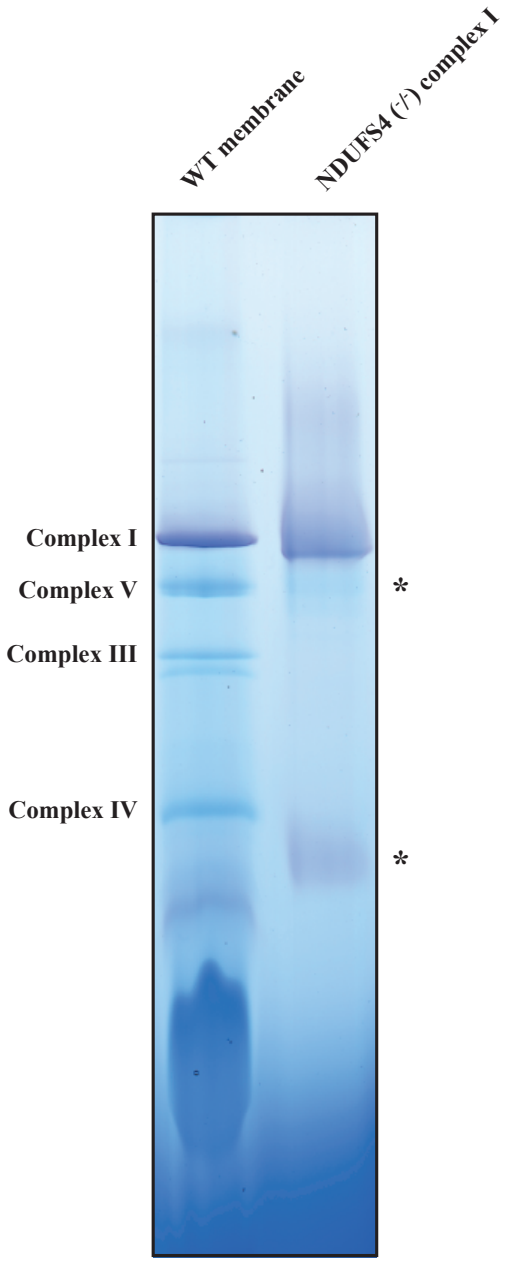


Figure 5.17: BN-PAGE of purified cross-linked *Ndufs4* ($^{-}/^{-}$) mouse complex I. Cross-linked *Ndufs4* ($^{-}/^{-}$) complex I (5 µg) and wild type mouse membranes (30 µg) were separated by BN-PAGE, and stained with 0.5 mg mL^{-1} NBT and 100 µM NADH in 20 mM Tris-HCl pH 7.55 at 32 °C. *Ndufs4* ($^{+}/^{+}$) membranes (WT membranes) and NDUFS4 ($^{-}/^{-}$) complex I. Asterisks indicate unknown proteins in the preparation.

5.13 Cryo-EM map of *Ndufs4* ($^{-/-}$) complex I from mouse kidney

The work presented here (with permission) on the cryo-EM map of *Ndufs4* ($^{-/-}$) complex I from mouse kidney was prepared by Zhan Yin. Grids were prepared from the purified sample according to Blaza *et al.,* (Blaza et al., 2018) that was then imaged at the UK National electron Bio-Imaging Centre (eBIC) at the Diamond Light Source. Images were processed according to the RELION pipeline procedure by Dr. Hannah Bridges (Fernandez-Leiro and Scheres, 2017).

The cryo-EM density map of *Ndufs4* ($^{-/-}$) complex I reached a global resolution of 8.4 Å according to the gold standard FSC = 0.143 criterion (Scheres and Chen, 2012). First, active mouse complex I was fitted into the density using the 'fit-in-map' function in Chimera (Pettersen et al., 2004a). The structure, which is immensely important in understanding how mutations in complex I result in neuromuscular disorders, is intact and its distinctive L-shape is similar to its wild type counterpart (Figure 5.18). There is an empty cavity in the structure (i.e. no density) where NDUFS4 is usually positioned. The relatively low resolution of the structure prevents observation of secondary structure features such helices and sheets. However, at 8.4 Å it is clear that there are no major shifts in the positions of the subunits, as they appear to fit relatively well into density.

Finally, there appears to be no unoccupied densities in the cryo-EM map that account for assembly factor(s), such as NDUFAF2 or ACAD9, with most of the empty densities representing the detergent belt. However, a higher resolution is structure is required to confirm this.

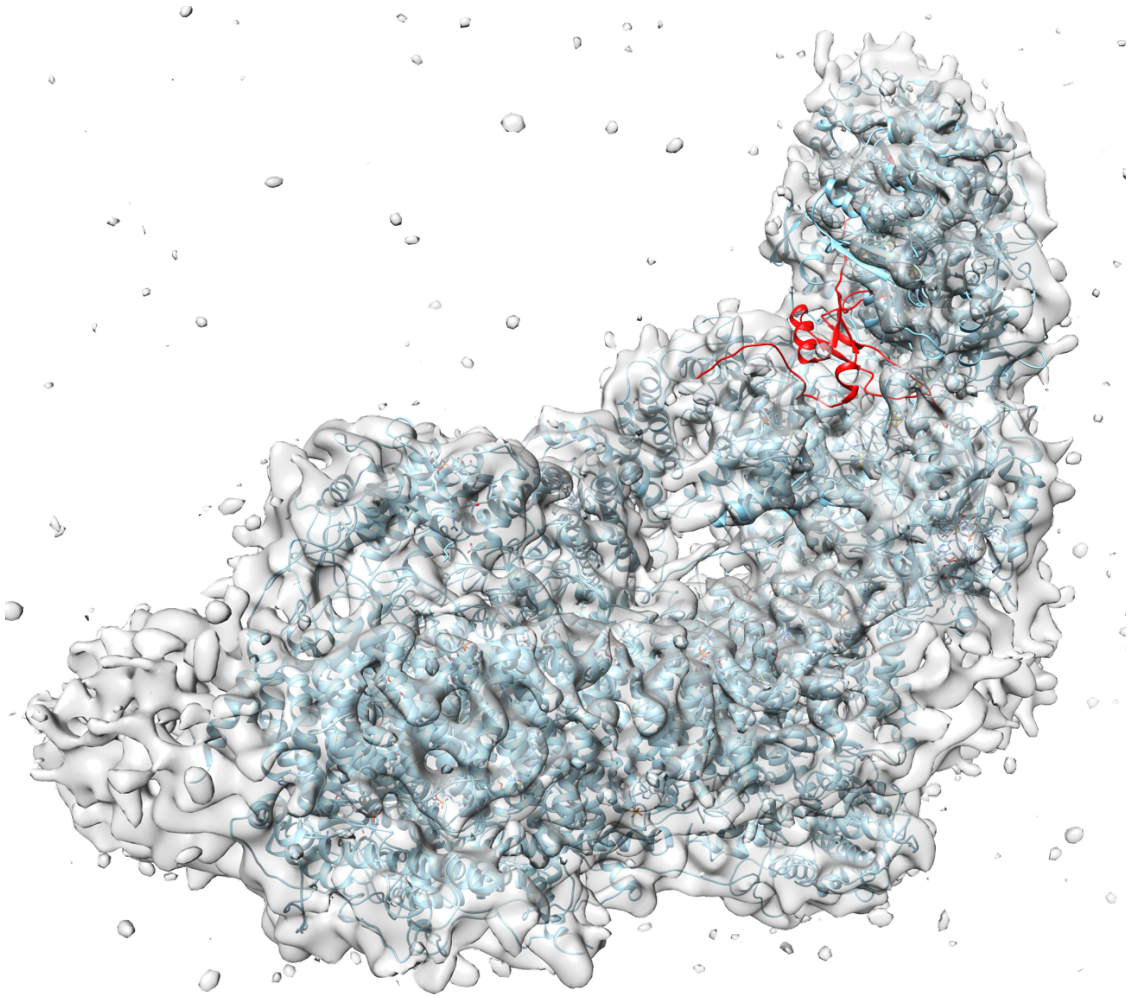


Figure 5.18: Cryo-EM map *Ndufs4* ($^{-}$/$^{-}$) complex I from mouse kidney. The structure of active mouse (secondary structure) fitted into the 8 Å density. NDUFS4 is shown in red while the rest of the protein is shown in blue.

5.14 Discussion

5.14.1 Summary

Deletion of *Ndufs4* gene in mice results in reduction in complex I content, and substantial loss of complex I catalytic activity in both membranes and purified preparation, consistent with what other groups have reported in different systems (Calvaruso et al., 2012; Kruse et al., 2008; Ugalde et al., 2004; Valsecchi et al., 2012). It was also established by mass spectrometry, cross-linking experiments and structural analysis that *Ndufs4* ($^{-/-}$) complex I is intact, but unstable, as observed during BN-PAGE analyses where complex I destabilised and breaks down into sub-complexes that were originally thought to be stalled complex I assembly intermediates (Valsecchi et al., 2012). Thus, it can be concluded that loss of NDUFS4 from complex I results in instability of the mature protein affecting its catalytic function. In order to stabilise *Ndufs4* ($^{-/-}$) complex I, a purification procedure was successfully established that ensures the enzyme is intact even under BN-PAGE conditions, where it would otherwise break apart.

5.14.2 *Ndufs4* intra-complex interaction and stability

It has been ascertained that *Ndufs4* ($^{-/-}$) complex I degradation is an artefact of BN-PAGE. Although NDUFS4 is a small protein (~15 kDa), it forms numerous interactions with other complex I subunits (Figure 5.19). 60% of its surface area is buried in interactions with other subunits of complex I (Table 4.3), the most amongst the supernumerary subunits alongside NDUFA1 (also 60% of its surface area is buried). In total, NDUFS4 interacts with 11 subunits, of which 6 are N-domain subunits (NDUFA12, NDUFS1 NDUFS6, NDUFV1, NDUFV2 & NDUFV3) and 5 are Q-domain subunits (NDUFA5, NDUFA6, NDUFA9, NDUFS3 and NDUFS8). Thus, the loss of NDUFS4 and its numerous interactions destabilises the complex, resulting in the loss of the N-domain (6 subunits of which are interacting with NDUFS4) and NDUFA7 of the Q-domain from complex I during BN-PAGE (Figure 5.5b). The 5 remaining subunits of complex I (NDUFA5, NDUFA6, NDUFA9, NDUFS3 and NDUFS8), which interact with NDUFS4, are not lost because they anchored more tightly to the rest of the complex (i.e. membrane domain). Nijtmans and co-workers reported that *Ndufs4* ($^{-/-}$) complex I is stabilised in the respiratory supercomplex I+III_2 (Calvaruso et al., 2012). However, complex I interactions with complex III and IV in supercomplexes are mediated by membrane-spanning

proteins (Gu et al., 2016; Guo et al., 2017; Letts et al., 2016a; Wu et al., 2016a), whilst NDUFS4 is in the hydrophilic fraction, therefore supercomplex mediated stabilisation is highly unlikely.

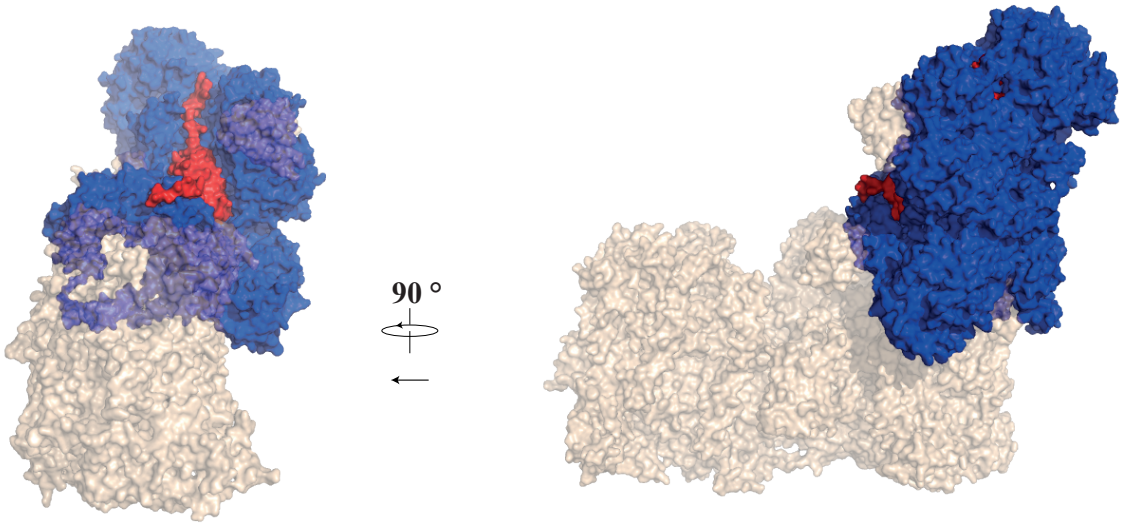


Figure 5.19: Visualisation of NDUFS4 interaction with other complex I subunits. A surface representation of the interaction of NDUFS4 (red) with complex I subunits (NDUFA12, NDUFS1 NDUFS6, NDUFV1, NDUFV2, NDUFV3, NDUFA5, NDUFA6, NDUFA9, NDUFS3 and NDUFS8) (blue), whilst remaining subunits are shown in wheat. The structure is rotated 90 ° about the shown axis. The structure is of the active mouse complex I presented in chapter 4.

5.14.3 Mechanistic dysfunction of *Ndufs4* ($^{-/-}$) mouse complex I

Catalytic efficiency of complex I has been reduced drastically as a result of deletion, in particular, it has been demonstrated that flavin-mediated NADH oxidation is rate limiting in the turnover of the enzyme. As there appears to be no major structural rearrangement of *Ndufs4* ($^{-/-}$) complex I, the changes may be the result of subtle changes in the structure or in the biophysical properties of the enzyme.

NDUFS4 spans almost the entire length of the electron transfer pathway of complex I and is in close proximity to all of the catalytically active iron-sulphur clusters and FMN (Figure 5.20). The protein is highly basic (predicted isoelectric point of 9.5), with most of its positive charge facing the iron-sulphur clusters (Figure 5.20b), enabling it to stabilise the reduced state of the clusters. It is known that the terminal cluster N2 in mammalian complex I is surrounded by three basic residues (Arg85, Arg105 and His190 in NDUFS2) that modulate its midpoint potential, in particular, Arg85, the absence of which in *T. thermophilus* complex I results in a lower midpoint potential for the adjacent cluster (less readily reduced) (Hirst and Roessler, 2016). Analogous to how the midpoint potential of cluster N2 is affected by nearby basic residues (Hirst and Roessler, 2016), it can be speculated that the loss of NDUFS4 may result in the lowering of the midpoint potentials of adjacent clusters, hindering electron transfer through the clusters. Moreover, Zickermann and co-workers observed an overall decrease in the signal amplitude of the EPR spectra of *ΔNdufs4 Y. lipolytica* complex I compared to complex I from wild (Kahlhöfer et al., 2017). In agreement with a possible role of NDUFS4 modulating potentials of the iron-sulphur clusters, there is a strong conservation of residues adjacent to the clusters (Figure 5.20c).

Although it is suggested here that electron transfer through the clusters is severely affected as a result of the loss of NDUFS4, more rigorous biochemical and biophysical, particularly EPR spectroscopy, to better define the clusters.

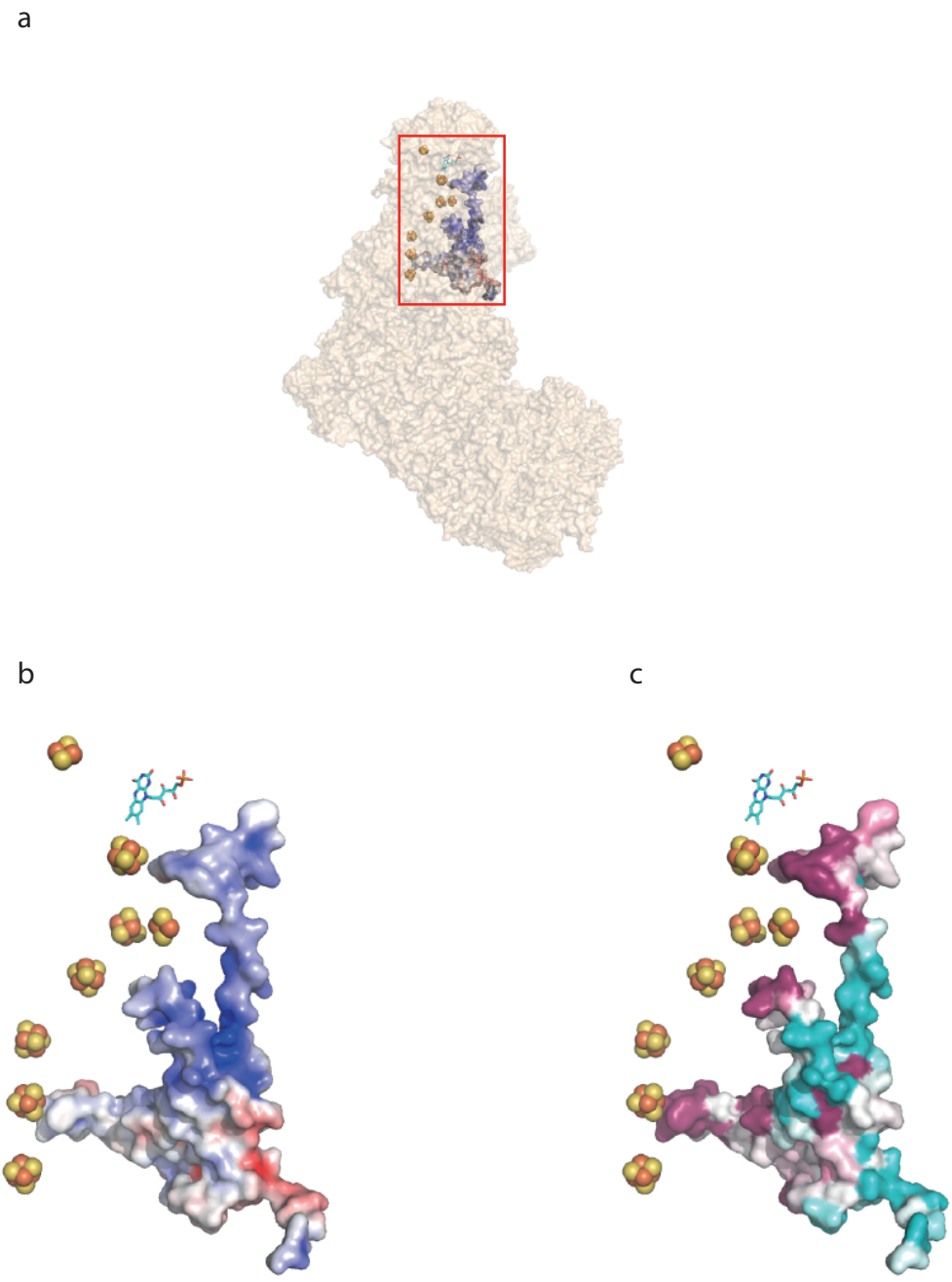


Figure 5.20: Visualisation of NDUFS4 Interaction with iron-sulphur clusters. (a) A surface representation of mouse complex I structure (wheat), with the iron-sulphur clusters (spheres), FMN and NDUFS4 highlighted in the red rectangle. (b) Electrostatic potential of NDUFS4 shown using the APBS plugin in Pymol with default parameters. Blue to white to red denotes basic, neutral and acidic residues, respectively. (c) NDUFS4 sequences from various species were aligned (using UNIREF90) and mapped onto the subunit using Consurf (Ashkenazy et al., 2010), with the default parameters. High, medium and low conservation are shown in purple, white and blue, respectively. The structure used is that of the active mouse complex I in chapter 4.

5.14.4 Future prospects

In order to identify subtle changes in the structure (i.e. altered confirmation compared to the wild type enzyme), the structure of *Ndufs4* ($^{-/-}$) complex I needs to be solved to a higher resolution using Cryo-EM. Ideally, less than 6 Å is required to observe secondary structure features such as sheets and helices, and ~4 Å or lower to resolve side chain densities. This will require stabilisation of the enzyme using cross-linkers, under similar conditions outlined in this chapter (see Chapter 5.12). For comparison, the structure of the wild type enzyme must be solved under the same conditions to ensure that the structure obtained is not an artefact of cross-linking.

Lastly, considering recent experiments demonstrating an improvement of phenotype when *Ndufs4* ($^{-/-}$) mice are exposed to chronic hypoxia, it would be interesting to study the effect of this treatment (Jain et al., 2016) on the complex I itself. It is not known whether complex I expression is adjusted, stability is affected or even modified under these conditions. This would require biochemical, biophysical and maybe even structural analyses to determine the molecular consequence of low oxygen tension on *Ndufs4* ($^{-/-}$) complex I.

Chapter 6: References

Abdrakhmanova, A., Zickermann, V., Bostina, M., Radermacher, M., Schägger, H., Kerscher, S., and Brandt, U. (2004). Subunit composition of mitochondrial complex I from the yeast *Yarrowia lipolytica*. Biochim. Biophys. Acta *1658*, 148–156.

Abdrakhmanova, A., Zwicker, K., Kerscher, S., Zickermann, V., and Brandt, U. (2006). Tight binding of NADPH to the 39-kDa subunit of complex I is not required for catalytic activity but stabilizes the multiprotein complex. Biochim. Biophys. Acta *1757*, 1676–1682.

Adrian, M., Dubochet, J., Lepault, J., and McDowall, A.W. (1984). Cryo-electron microscopy of viruses. Nature *308*, 32–36.

Afonine, P. V., Headd, J.J., Terwilliger, T.C., and Adams, P.D. (2013). New tool : phenix . real _ space _ refine. Comput. Crystallogr. Newsl. *4*, 43–44.

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., et al. (1981). Sequence and organization of the human mitochondrial genome. Nature *290*, 457–465.

Anderson, S.L., Chung, W.K., Frezzo, J., Papp, J.C., Ekstein, J., DiMauro, S., and Rubin, B.Y. (2008). A novel mutation in NDUFS4 causes Leigh syndrome in an Ashkenazi Jewish family. J. Inherit. Metab. Dis. *31*, 461–467.

Andrews, B., Carroll, J., Ding, S., Fearnley, I.M., and Walker, J.E. (2013). Assembly factors for the membrane arm of human complex I. Proc. Natl. Acad. Sci. U. S. A. *110*, 18934–18939.

Antonicka, H., Ogilvie, I., Taivassalo, T., Anitori, R.P., Haller, R.G., Vissing, J., Kennaway, N.G., and Shoubridge, E. a. (2003). Identification and Characterization of a Common Set of Complex I Assembly Intermediates in Mitochondria from Patients with Complex I Deficiency. J. Biol. Chem. *278*, 43081–43088.

Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E.,

Huang, H., Lopez, R., Magrane, M., et al. (2017). UniProt: the universal protein knowledgebase. Nucleic Acids Res. *45*, D158–D169.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. *38*, 529–533.

Baradaran, R., Berrisford, J.M., Minhas, G.S., and Sazanov, L. a (2013). Crystal structure of the entire respiratory complex I. Nature *494*, 443–448.

Bartesaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J.L.S., and Subramaniam, S. (2015). 2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor. Science *348*, 1147–1151.

Baughman, J.M., Perocchi, F., Girgis, H.S., Plovanich, M., Belcher-Timme, C.A., Sancak, Y., Bao, X.R., Strittmatter, L., Goldberger, O., Bogorad, R.L., et al. (2011). Integrative genomics identifies MCU as an essential component of the mitochondrial calcium uniporter. Nature *476*, 341–345.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The protein data bank. Nucleic Acids Res. *28*, 235–242.

Birrell, J., and Hirst, J. (2013). Investigation of NADH Binding, Hydride Transfer, and NAD+ Dissociation during NADH Oxidation by Mitochondrial Complex I Using Modified Nicotinamide. Biochemistry *52*, 4048–4055.

Birrell, J.A., and Hirst, J. (2010). Truncation of subunit ND2 disrupts the threefold symmetry of the antiporter-like subunits in complex I from higher metazoans. *584*, 4247–4252.

Birrell, J. a, King, M.S., and Hirst, J. (2011). A ternary mechanism for NADH oxidation by positively charged electron acceptors, catalyzed at the flavin site in respiratory complex I. FEBS Lett. *585*, 2318–2322.

Birrell, J.A., Yakovlev, G., and Hirst, J. (2009). Reactions of the flavin mononucleotide in

complex I: a combined mechanism describes NADH oxidation coupled to the reduction of APAD+, ferricyanide, or molecular oxygen. Biochemistry *48*, 12005–12013.

Blanchi, C., Genova, M.L., Castelli, G.P., and Lenaz, G. (2004). The mitochondrial respiratory chain is partially organized in a supercomplex assembly: Kinetic evidence using flux control analysis. J. Biol. Chem. *279*, 36562–36569.

Blaza, J.N., Serreli, R., Jones, A.J.Y., Mohammed, K., and Hirst, J. (2014). Kinetic evidence against partitioning of the ubiquinone pool and the catalytic relevance of respiratory-chain supercomplexes. Proc. Natl. Acad. Sci. *111*, 15735–15740.

Blaza, J.N., Vinothkumar, K.R., and Hirst, J. (2018). Structure of the Deactive State of Mammalian Respiratory Complex I. Structure *26*, 312–319.

Bologna, F.P., Andreo, C.S., and Drincovich, M.F. (2007). *Escherichia coli* malic enzymes: Two isoforms with substantial differences in kinetic properties, metabolic regulation, and structure. J. Bacteriol. *189*, 5937–5946.

Boxma, B., De Graaf, R.M., Van Der Staay, G.W.M., Van Alen, T.A., Ricard, G., Gabaldón, T., Van Hoek, A.H.A.M., Moon-Van Der Staay, S.Y., Koopman, W.J.H., Van Hellemond, J.J., et al. (2005). An anaerobic mitochondrion that produces hydrogen. Nature *434*, 74–79.

Braymer, J.J., and Lill, R. (2017). Iron–sulfur cluster biogenesis and trafficking in mitochondria. J. Biol. Chem. *292*, 12754–12763.

Le Breton, N., Wright, J.J., Jones, A.J.Y., Salvadori, E., Bridges, H.R., Hirst, J., and Roessler, M.M. (2017). Using Hyperfine Electron Paramagnetic Resonance Spectroscopy to Define the Proton-Coupled Electron Transfer Reaction at Fe–S Cluster N2 in Respiratory Complex I. J. Am. Chem. Soc. *139*, 16319–16326.

Bridges, H.R., Bill, E., and Hirst, J. (2012). Mössbauer spectroscopy on respiratory complex I: the iron-sulfur cluster ensemble in the NADH-reduced enzyme is partially oxidized. Biochemistry *51*, 149–158.

Bridges, H.R., Mohammed, K., Harbour, M.E., and Hirst, J. (2017). Subunit NDUFV3 is present in two distinct isoforms in mammalian complex I. Biochim. Biophys. Acta *1858*, 197–207.

Buchanan, S.K., and Walker, J.E. (1996). Large-scale chromatographic purification of F1 Fo - ATPase and complex I from bovine heart mitochondria. Biochem. J. *349*, 343–349.

Budde, S.M.S., van den Heuvel, L.P.W.J., Smeets, R.J.P., Skladal, D., Mayr, J.A., Boelen, C., Petruzzella, V., Papa, S., and Smeitink, J.A.M. (2003). Clinical heterogeneity in patients with mutations in the NDUFS4 gene of mitochondrial complex I. J. Inherit. Metab. Dis. *26*, 813–815.

Calvaruso, M.A., Willems, P., Van den brand, M., Valsecchi, F., Kruse, S., Palmiter, R., Smeitink, J., and Nijtmans, L. (2012). Mitochondrial complex III stabilizes complex I in the absence of NDUFS4 to provide partial activity. Hum. Mol. Genet. *21*, 115–120.

Calvo, S.E., Tucker, E.J., Compton, A.G., Kirby, D.M., Crawford, G., Burtt, N.P., Rivas, M., Guiducci, C., Bruno, D.L., Goldberger, O.A., et al. (2010). High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. Nat. Genet. *42*, 851–858.

Carroll, J., Fearnley, I.M., Shannon, R.J., Hirst, J., and Walker, J.E. (2003). Analysis of the subunit composition of complex I from bovine heart mitochondria. Mol. Cell. Proteomics *2*, 117–126.

Carroll, J., Fearnley, I.M., Wang, Q., and Walker, J.E. (2009). Measurement of the molecular masses of hydrophilic and hydrophobic subunits of ATP synthase and complex I in a single experiment. Anal. Biochem. *395*, 249–255.

Carroll, J., Ding, S., Fearnley, I.M., and Walker, J.E. (2013). Post-translational modifications near the quinone binding site of mammalian complex I. J. Biol. Chem. *288*, 24799–24808.

Chae, P.S., Rasmussen, S.G.F., Rana, R.R., Gotfryd, K., Chandra, R., Goren, M.A., Kruse, A.C., Nurva, S., Loland, C.J., Pierre, Y., et al. (2010). Maltose–neopentyl glycol (MNG)

amphiphiles for solubilization, stabilization and crystallization of membrane proteins. Nat. Methods *7*, 1003–1008.

Chouchani, E.T., Methner, C., Nadtochiy, S.M., Logan, A., Pell, V.R., Ding, S., James, A.M., Cochemé, H.M., Reinhold, J., Lilley, K.S., et al. (2013). Cardioprotection by S-nitrosation of a cysteine switch on mitochondrial complex I. Nat. Med. *19*, 753–759.

Chouchani, E.T., Pell, V.R., Gaude, E., Aksentijević, D., Sundier, S.Y., Robb, E.L., Logan, A., Nadtochiy, S.M., Ord, E.N.J., Smith, A.C., et al. (2014). Ischaemic accumulation of succinate controls reperfusion injury through mitochondrial ROS. Nature *515*, 431–435.

Choudhry, H., and Harris, A.L. (2018). Advances in Hypoxia-Inducible Factor Biology. Cell Metab. *27*, 281–298.

Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., et al. (2012). CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Comput. Biol. *8*, e1002708.

Efremov, R.G., and Sazanov, L. a (2011). Structure of the membrane domain of respiratory complex I. Nature *476*, 414–420.

Elurbe, D.M., and Huynen, M.A. (2016). The origin of the supernumerary subunits and assembly factors of complex I: A treasure trove of pathway evolution. Biochim. Biophys. Acta - Bioenerg. *1857*, 971–979.

Embley, T.M., and Martin, W. (2006). Eukaryotic evolution, changes and challenges. Nature.
Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 486–501.

Eriksson, S., Munch-Petersen, B., Johansson, K., and Ecklund, H. (2002). Structure and function of cellular deoxyribonucleoside kinases. Cell. Mol. Life Sci. *59*, 1327–1346.

Fassone, E., and Rahman, S. (2012). Complex I deficiency: clinical features, biochemistry and

molecular genetics. J. Med. Genet. *49*, 578–590.

Fassone, E., Taanman, J.-W., Hargreaves, I.P., Sebire, N.J., Cleary, M. A, Burch, M., and Rahman, S. (2011). Mutations in the mitochondrial complex I assembly factor NDUFAF1 cause fatal infantile hypertrophic cardiomyopathy. J. Med. Genet. *48*, 691–697.

Fedor, J.G., Jones, A.J.Y., Di Luca, A., Kaila, V.R.I., and Hirst, J. (2017). Correlating kinetic and structural data on ubiquinone binding and reduction by respiratory complex I. Proc. Natl. Acad. Sci. *114*, 12737–12742.

Fernandez-Leiro, R., and Scheres, S.H.W. (2016). Unravelling biological macromolecules with cryo-electron microscopy. Nature *537*, 339–346.

Fernandez-Leiro, R., and Scheres, S.H.W. (2017). A pipeline approach to single-particle processing in RELION. In Acta Crystallographica Section D: Structural Biology, pp. 496–502.

Fernández-vizarra, E., Ferrín, G., Pérez-martos, A., Fernández-silva, P., Zeviani, M., and Antonio, J. (2010). Mitochondrion Isolation of mitochondria for biogenetical studies : An update. Mitochondrion *10*, 253–262.

Fernández-Vizarra, E., López-Pérez, M.J., and Enriquez, J.A. (2002). Isolation of biogenetically competent mitochondria from mammalian tissues and cultured cells. Methods *26*, 292–297.

Ferrari, M., Jain, I.H., Goldberger, O., Rezoagli, E., Thoonen, R., Chen, K.-H., Sosnovik, D.E., Scherrer-Crosbie, M., Mootha, V.K., and Zapol, W.M. (2017). Hypoxia treatment reverses neurodegenerative disease in a mouse model of Leigh syndrome. Proc. Natl. Acad. Sci. U. S. A. *114*, E4241–E4250.

Fiedorczuk, K., Letts, J.A., Degliesposti, G., Kaszuba, K., Skehel, M., and Sazanov, L.A. (2016). Atomic structure of the entire mammalian mitochondrial complex I. Nature *538*, 406–410.

Forneris, F., Orru, R., Bonivento, D., Chiarelli, L.R., and Mattevi, A. (2009). Thermo FAD, a

Thermofluor® -adapted flavin ad hoc detection system for protein folding and ligand binding. FEBS J. *276*, 2833–2840.

Frank, J. (1975). Averaging of low exposure electron micrographs of non-periodic objects. Ultramicroscopy *1*, 159–162.

Freya, T.G., and Mannellab, C.A. (2000). The internal structure of mitochondria. Trends Biochem. Sci. *25*, 319–324.

Galkin, A., Meyer, B., Wittig, I., Karas, M., Schägger, H., Vinogradov, A., and Brandt, U. (2008). Identification of the mitochondrial ND3 subunit as a structural component involved in the active/deactive enzyme transition of respiratory complex I. J. Biol. Chem. *283*, 20907–20913.

Gamiz-Hernandez, A.P., Jussupow, A., Johansson, M.P., and Kaila, V.R.I. (2017). Terminal Electron-Proton Transfer Dynamics in the Quinone Reduction of Respiratory Complex i. J. Am. Chem. Soc. *139*, 16282–16288.

Glaeser, R.M. (2016). Specimen Behavior in the Electron Beam. Methods Enzymol.

Gu, J., Wu, M., Guo, R., Yan, K., Lei, J., Gao, N., and Yang, M. (2016). The architecture of the mammalian respirasome. Nature *537*, 639–643.

Guénebaut, V., Vincentelli, R., Mills, D., Weiss, H., and Leonard, K.R. (1997). Three-dimensional structure of NADH-dehydrogenase from Neurospora crassa by electron microscopy and conical tilt reconstruction. J. Mol. Biol. *265*, 409–418.

Guerrero-Castillo, S., Baertling, F., Kownatzki, D., Wessels, H.J., Arnold, S., Brandt, U., and Nijtmans, L. (2016). The Assembly Pathway of Mitochondrial Respiratory Chain Complex I. Cell Metab. *25*, 128–139.

Guo, R., Zong, S., Wu, M., Gu, J., and Yang, M. (2017). Architecture of Human Mitochondrial Respiratory Megacomplex $I_2III_2IV_2$. Cell *170*, 1247–1257.

Hatefi, Y., Haavik, A.G., and Griffiths, D.E. (1962). Studies on the electron transfer system.

XL. Preparation and properties of mitochondrial DPNH-coenzyme Q reductase. J. Biol. Chem. *237*, 1676–1680.

Heide, H., Bleier, L., Steger, M., Ackermann, J., Dröse, S., Schwamb, B., Zörnig, M., Reichert, A.S., Koch, I., Wittig, I., et al. (2012). Complexome Profiling Identifies TMEM126B as a Component of the Mitochondrial Complex I Assembly Complex. Cell Metab. *16*, 538–549.

Henderson, R. (1990). Cryo-Protection of Protein Crystals against Radiation Damage in Electron and X-Ray Diffraction. Proc. R. Soc. B Biol. Sci. *241*, 6–8.

van den Heuvel, L., Ruitenbeek, W., Smeets, R., Gelman-Kohan, Z., Elpeleg, O., Loeffen, J., Trijbels, F., Mariman, E., de Bruijn, D., and Smeitink, J. (1998). Demonstration of a new pathogenic mutation in human complex I deficiency: a 5-bp duplication in the nuclear gene encoding the 18-kD (AQDQ) subunit. Am. J. Hum. Genet. *62*, 262–268.

Hiltunen, J.K., Chen, Z., Haapalainen, A.M., Wierenga, R.K., and Kastaniotis, A.J. (2010). Mitochondrial fatty acid synthesis--an adopted set of enzymes making a pathway of major importance for the cellular metabolism. Prog. Lipid Res. *49*, 27–45.

Hinchliffe, P., and Sazanov, L.A. (2005). Organization of iron-sulfur clusters in respiratory complex I. Science *309*, 771–774.

Hirst, J., and Roessler, M.M. (2016). Energy conversion, redox catalysis and generation of reactive oxygen species by respiratory complex I. Biochim. Biophys. Acta - Bioenerg. *1857*, 872–883.

Hirst, J., Carroll, J., Fearnley, I.M., Shannon, R.J., and Walker, J.E. (2003). The nuclear encoded subunits of complex I from bovine heart mitochondria. Biochim. Biophys. Acta - Bioenerg. *1604*, 135–150.

Hofhaus, G., Weiss, H., and Leonard, K. (1991). Electron microscopic analysis of the peripheral and membrane parts of mitochondrial NADH dehydrogenase (Complex I). J. Mol. Biol. *221*, 1027–1043.

Hryc, C.F., Chen, D.-H., Afonine, P. V., Jakana, J., Wang, Z., Haase-Pettingell, C., Jiang, W., Adams, P.D., King, J.A., Schmid, M.F., et al. (2017). Accurate model annotation of a near-atomic resolution cryo-EM map. Proc. Natl. Acad. Sci. *114*, 3103–3108.

Hu, H., Nan, J., Sun, Y., Zhu, D., Xiao, C., Wang, Y., Zhu, L., Wu, Y., Zhao, J., Wu, R., et al. (2017). Electron leak from NDUFA13 within mitochondrial complex I attenuates ischemia-reperfusion injury via dimerized STAT3. Proc. Natl. Acad. Sci. U. S. A. *114*, 11908–11913.

Hunte, C. (2005). Specific protein-lipid interactions in membrane proteins. Biochem. Soc. Trans. *33*, 938–942.

Hunte, C., Zickermann, V., and Brandt, U. (2010). Functional modules and structural basis of conformational coupling in mitochondrial complex I. Science *329*, 448–451.

Jain, I.H., Zazzeron, L., Goli, R., Alexa, K., Schatzman-Bone, S., Dhillon, H., Goldberger, O., Peng, J., Shalem, O., Sanjana, N.E., et al. (2016). Hypoxia as a therapy for mitochondrial disease. Science *352*, 54–61.

Johnson, S.C., Yanos, M.E., Kayser, E.-B., Quintana, A., Sangesland, M., Castanza, A., Uhde, L., Hui, J., Wall, V.Z., Gagnidze, A., et al. (2013). mTOR inhibition alleviates mitochondrial disease in a mouse model of Leigh syndrome. Science *342*, 1524–1528.

Jones, A.J.Y., and Hirst, J. (2013). A spectrophotometric coupled enzyme assay to measure the activity of succinate dehydrogenase. Anal. Biochem. *442*, 19–23.

Jones, A.J.Y., Blaza, J.N., Bridges, H.R., May, B., Moore, A.L., and Hirst, J. (2016). A Self-Assembled Respiratory Chain that Catalyzes NADH Oxidation by Ubiquinone-10 Cycling between Complex I and the Alternative Oxidase. Angew. Chemie Int. Ed. *55*, 728–731.

Jones, A.J.Y., Blaza, J.N., Varghese, F., and Hirst, J. (2017). Respiratory Complex I in *Bos taurus* and Paracoccus denitrificans Pumps Four Protons across the Membrane for Every NADH Oxidized. J. Biol. Chem. *292*, 4987–4995.

Kahlhöfer, F., Kmita, K., Wittig, I., Zwicker, K., and Zickermann, V. (2017). Accessory subunit NUYM (NDUFS4) is required for stability of the electron input module and activity of mitochondrial complex I. Biochim. Biophys. Acta - Bioenerg. *1858*, 175–181.

Karnkowska, A., Vacek, V., Zubáčová, Z., Treitli, S.C., Petrželková, R., Eme, L., Novák, L., Žárský, V., Barlow, L.D., Herman, E.K., et al. (2016). A Eukaryote without a Mitochondrial Organelle. Curr. Biol. *26*, 1274–1284.

Ke, B.-X., Pepe, S., Grubb, D.R., Komen, J.C., Laskowski, A., Rodda, F.A., Hardman, B.M., Pitt, J.J., Ryan, M.T., Lazarou, M., et al. (2012). Tissue-specific splicing of an Ndufs6 gene-trap insertion generates a mitochondrial complex I deficiency-specific cardiomyopathy. Proc. Natl. Acad. Sci. U. S. A. *109*, 6165–6170.

Kimanius, D., Forsberg, B.O., Scheres, S.H.W., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife *5*, e18722.

Kotlyar, A.B., and Vinogradov, A.D. (1990). Slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase. Biochim. Biophys. Acta *1019*, 151–158.

Krissinel, E., and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. *372*, 774–797.

Kruse, S.E., Watt, W.C., Marcinek, D.J., Kapur, R.P., Schenkman, K. a., and Palmiter, R.D. (2008). Mice with Mitochondrial Complex I Deficiency Develop a Fatal Encephalomyopathy. Cell Metab. *7*, 312–320.

Kühlbrandt, W. (2015). Structure and function of mitochondrial membrane protein complexes. BMC Biol. *13*, 89.

Kunji, E.R.S., and Robinson, A.J. (2010). Coupling of proton and substrate translocation in the transport cycle of mitochondrial carriers. Curr. Opin. Struct. Biol. *20*, 440–447.

Kussmaul, L., and Hirst, J. (2006). The mechanism of superoxide production by

NADH:ubiquinone oxidoreductase (complex I) from bovine heart mitochondria. Proc. Natl. Acad. Sci. U. S. A. *103*, 7607–7612.

Lapuente-Brun, E., Moreno-Loshuertos, R., Acín-Pérez, R., Latorre-Pellicer, A., Colás, C., Balsa, E., Perales-Clemente, E., Quirós, P.M., Calvo, E., Rodríguez-Hernández, M. a, et al. (2013). Supercomplex assembly determines electron flux in the mitochondrial electron transport chain. Science *340*, 1567–1570.

Letts, J.A., Fiedorczuk, K., and Sazanov, L.A. (2016a). The architecture of respiratory supercomplexes. Nature *537*, 644–648.

Letts, J.A., Degliesposti, G., Fiedorczuk, K., Skehel, M., and Sazanov, L.A. (2016b). Purification of ovine respiratory complex I results in a highly active and stable preparation. J. Biol. Chem. *291*, 24657–24675.

Lin, C.S., Sharpley, M.S., Fan, W., Waymire, K.G., Sadun, A.A., Carelli, V., Ross-Cisneros, F.N., Baciu, P., Sung, E., McManus, M.J., et al. (2012). Mouse mtDNA mutant model of Leber hereditary optic neuropathy. Proc. Natl. Acad. Sci. *109*, 20065–20070.

Di Luca, A., and Kaila, V.R.I. (2018). Global collective motions in the mammalian and bacterial respiratory complex I. Biochim. Biophys. Acta - Bioenerg. *1859*, 326–332.

Maklashina, E., Kotlyar, A.B., and Cecchini, G. (2003). Active/de-active transition of respiratory complex I in bacteria, fungi, and animals. Biochim. Biophys. Acta - Bioenerg. *1606*, 95–103.

Martín, M. a, Blázquez, A., Gutierrez-Solana, L.G., Fernández-Moreira, D., Briones, P., Andreu, A.L., Garesse, R., Campos, Y., and Arenas, J. (2005). Leigh syndrome associated with mitochondrial complex I deficiency due to a novel mutation in the NDUFS1 gene. Arch. Neurol. *62*, 659–661.

Martinou, J.C., and Youle, R.J. (2011). Mitochondria in Apoptosis: Bcl-2 Family Members and Mitochondrial Dynamics. Dev. Cell *21*, 92–101.

Mathiesen, C., and Hägerhäll, C. (2002). Transmembrane topology of the NuoL, M and N

subunits of NADH:quinone oxidoreductase and their homologues among membrane-bound hydrogenases and bona fide antiporters. Biochim. Biophys. Acta *1556*, 121–132.

Mckenzie, M., and Ryan, M.T. (2010). Assembly factors of human mitochondrial complex I and their defects in disease. IUBMB Life *62*, 497–502.

McMullan, G., Faruqi, A.R., Clare, D., and Henderson, R. (2014). Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. Ultramicroscopy *147*, 156–163.

McMullan, G., Faruqi, A.R., and Henderson, R. (2016). Direct Electron Detectors. Methods Enzymol. *579*, 1–17.

Di Meo, I., Marchet, S., Lamperti, C., Zeviani, M., and Viscomi, C. (2017). AAV9-based gene therapy partially ameliorates the clinical phenotype of a mouse model of Leigh syndrome. Gene Ther. *24*, 661–667.

Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M.I., Pragani, R., Boxer, M.B., Earl, L.A., Milne, J.L.S., et al. (2016). Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. Cell *165*, 1698–1707.

Meyer, E.H., Solheim, C., Tanz, S.K., Bonnard, G., and Millar, A.H. (2011). Insights into the composition and assembly of the membrane arm of plant complex I through analysis of subcomplexes in Arabidopsis mutant lines. J. Biol. Chem. *266*, 26081-26092

Meyerson, J.R., Rao, P., Kumar, J., Chittori, S., Banerjee, S., Pierson, J., Mayer, M.L., and Subramaniam, S. (2015). Self-assembled monolayers improve protein distribution on holey carbon cryo-EM supports. Sci. Rep. *4*, 7084.

Mikkelsen, N.E., Johansson, K., Karlsson, A., Knecht, W., Andersen, G., Piškur, J., Munch-Petersen, B., and Eklund, H. (2003). Structural basis for feedback inhibition of the deoxyribonucleoside salvage pathway: Studies of the drosophila deoxyribonucleoside kinase. Biochemistry *42*, 5706–5712.

Milenkovic, D., Blaza, J.N., Larsson, N.G., and Hirst, J. (2017). The Enigma of the Respiratory Chain Supercomplex. Cell Metab. *25*, 765–776.

Minakami, S., Schindler, F.J., and Estabrook, R.W. (1964). Hydrogen Transfer Between Reduced Diphosphopyridine Nucleotide Dehydrogenase And The Respiratory Chain. II. An Initial Lag In The Oxidation Of Reduced Diphosphopyridine Nucleotide. J. Biol. Chem. *239*, 2049–2054.

Mohammed, K. (2015). Purification and characterisation of respiratory complex I (NADH:ubiquinone oxidoreductase) from five tissues of *Rattus norvegicus*. PhD thesis. Cambridge University.

Mourier, T., Hansen, A.J., Willerslev, E., and Arctander, P. (2001). Letter to the Editor The Human Genome Project Reveals a Continuous Transfer of Large Mitochondrial Fragments to the Nucleus. Mol. Biol. Evol *18*, 1833–1837.

Nakamoto, R.K., Baylis Scanlon, J.A., and Al-Shawi, M.K. (2008). The rotary mechanism of the ATP synthase. Arch. Biochem. Biophys. *476*, 43–50.

Nass, S., and Nass, M.M. (1963). Intramitochondrial fibers with DNA characteristics. II. Enzymatic and other hydrolytic treatments. J. Cell Biol. *19*, 613–629.

Ogilvie, I., Kennaway, N.G., and Shoubridge, E.A. (2005). A molecular chaperone for mitochondrial complex I assembly is mutated in a progressive encephalopathy. J. Clin. Invest. *115*, 2784–2792.

Orlova, E. V., and Saibil, H.R. (2011). Structural analysis of macromolecular assemblies by electron microscopy. Chem. Rev. *111*, 7710–7748.

Ortigoza-Escobar, J.D., Oyarzabal, A., Montero, R., Artuch, R., Jou, C., Jiménez, C., Gort, L., Briones, P., Muchart, J., López-Gallardo, E., et al. (2016). Ndufs4 related Leigh syndrome: A case report and review of the literature. Mitochondrion *28*, 73–78.

Page, C.C., Moser, C.C., and Dutton, P.L. (2003). Mechanism for electron transfer within and between proteins. Curr. Opin. Chem. Biol. *7*, 551–556.

Papa, S., Sardanelli, A.M., Cocco, T., Speranza, F., Scacco, S.C., and Technikova-Dobrova, Z. (1996). The nuclear-encoded 18 kDa (IP) AQDQ subunit of bovine heart complex I is phosphorylated by the mitochondrial cAMP-dependent protein kinase. FEBS Lett. *379*, 299–301.

Peralta, S., Torraco, A., Wenz, T., Garcia, S., Diaz, F., and Moraes, C.T. (2014). Partial complex I deficiency due to the CNS conditional ablation of Ndufa5 results in a mild chronic encephalopathy but no increase in oxidative damage. Hum. Mol. Genet. *23*, 1399–1412.

Petrek, M., Otyepka, M., Banás, P., Kosinová, P., Koca, J., and Damborský, J. (2006). CAVER: a new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinformatics *7*, 316.

Petruzzella, V., Vergari, R., Puzziferri, I., Boffoli, D., Lamantea, E., Zeviani, M., and Papa, S. (2001). A nonsense mutation in the NDUFS4 gene encoding the 18 kDa (AQDQ) subunit of complex I abolishes assembly and activity of the complex in a patient with Leigh-like syndrome. Hum. Mol. Genet. *10*, 529–535.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—A Visualization System for Exploratory Research and Analysis. J Comput Chem *25*, 1605–1612.

Pryde, K.R., and Hirst, J. (2011). Superoxide is produced by the reduced flavin in mitochondrial complex I: a single, unified mechanism that applies during both forward and reverse electron transfer. J. Biol. Chem. *286*, 18056–18065.

Pullman, M.E., Penefsky, H.S., Datta, A., and Racker, E. (1960). Partial resolution of the enzymes catalyzing oxidative phosphorylation. I. Purification and properties of soluble dinitrophenol-stimulated adenosine triphosphatase. J. Biol. Chem. *235*, 3322–3329.

Quintana, A., Kruse, S.E., Kapur, R.P., Sanz, E., and Palmiter, R.D. (2010). Complex I deficiency due to loss of Ndufs4 in the brain results in progressive encephalopathy resembling Leigh syndrome. Proc. Natl. Acad. Sci. U. S. A. *107*, 10996–11001.

Reda, T., Barker, C.D., and Hirst, J. (2008). Reduction of the iron-sulfur clusters in mitochondrial NADH:ubiquinone oxidoreductase (complex I) by EuII-DTPA, a very low potential reductant. Biochemistry *47*, 8885–8893.

Russo, C.J., and Passmore, L.A. (2014). Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. Science *346*, 1377–1380.

Sabini, E., Hazra, S., Ort, S., Konrad, M., and Lavie, A. (2008). Structural Basis for Substrate Promiscuity of dCK. J. Mol. Biol. *378*, 607–621.

Sato, M., and Sato, K. (2011). Degradation of paternal mitochondria by fertilization-triggered autophagy in C. elegans embryos. Science *334*, 1141–1144.

Sazanov, L.A. (2015). A giant molecular proton pump: Structure and mechanism of respiratory complex I. Nat. Rev. Mol. Cell Biol. *16*, 375–388.

Sazanov, L.A., and Hinchliffe, P. (2006). Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*. Science *311*, 1430–1436.

Sazanov, L.A., and Walker, J.E. (2000). Cryo-electron Crystallography of Two Sub-complexes of Bovine Complex I Reveals the Relationship between the Membrane and Peripheral Arms.

Sazanov, L.A., Peak-chew, S.Y., Fearnley, I.M., and Walker, J.E. (2000). Resolution of the Membrane Domain of Bovine Complex I into Subcomplexes : Implications for the Structural Organization of the Enzyme. Biochemistry *39*, 7229–7235.

Scacco, S., Vergari, R., Scarpulla, R.C., Technikova-Dobrova, Z., Sardanelli, A., Lambo, R., Lorusso, V., and Papa, S. (2000). cAMP-dependent phosphorylation of the nuclear encoded 18-kDa (IP) subunit of respiratory complex I and activation of the complex in serum- starved mouse fibroblast cultures. J. Biol. Chem. *275*, 17578–17582.

Scacco, S., Petruzzella, V., Budde, S., Vergari, R., Tamborra, R., Panelli, D., Van den Heuvel, L.P., Smeitink, J.A., and Papa, S. (2003). Pathological Mutations of the Human NDUFS4 Gene of the 18-kDa (AQDQ) Subunit of Complex I Affect the Expression of the Protein and the Assembly and Function of the Complex. J. Biol. Chem. *278*, 44161–44167.

Schägger, H., and von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal. Biochem. *199*, 223–231.

Schägger, H., and Pfeiffer, K. (2000). Supercomplexes in the respiratory chains of yeast and mammalian mitochondria. EMBO J. *19*, 1777–1783.

Scheres, S.H.W. (2014). Beam-induced motion correction for sub-megadalton cryo-EM particles. Elife *3*, e03665.

Scheres, S.H.W. (2016). Processing of Structurally Heterogeneous Cryo-EM Data in RELION. Methods Enzymol. *579*, 125–157.

Scheres, S.H.W., and Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. Nat. Methods *9*, 853–854.

Schilling, B., Aggeler, R., Schulenberg, B., Murray, J., Row, R.H., Capaldi, R.A., and Gibson, B.W. (2005). Mass spectrometric identification of a novel phosphorylation site in subunit NDUFA10 of bovine mitochondrial complex I. FEBS Lett. *579*, 2485–2490.

Sharma, V., Belevich, G., Gamiz-Hernandez, A.P., Róg, T., Vattulainen, I., Verkhovskaya, M.L., Wikström, M., Hummer, G., and Kaila, V.R.I. (2015). Redox-induced activation of the proton pump in the respiratory complex I. Proc. Natl. Acad. Sci. *112*, 11571–11576.

Sharpley, M.S., Shannon, R.J., Draghi, F., and Hirst, J. (2006). Interactions between phospholipids and NADH:ubiquinone oxidoreductase (complex I) from bovine mitochondria. Biochemistry *45*, 241–248.

Shi, J.-M., Pei, J., Liu, E.-Q., and Zhang, L. (2017). Bis(sulfosuccinimidyl) suberate (BS3)

crosslinking analysis of the behavior of amyloid-β peptide in solution and in phospholipid membranes. PLoS One *12*, e0173871.

Stanke, M., and Morgenstern, B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res. *33*, W465-467.

De Stefani, D., Raffaello, A., Teardo, E., Szabó, I., and Rizzuto, R. (2011). A forty-kilodalton protein of the inner membrane is the mitochondrial calcium uniporter. Nature *476*, 336–340.

Stroud, D.A., Surgenor, E.E., Formosa, L.E., Reljic, B., Frazier, A.E., Dibley, M.G., Osellame, L.D., Stait, T., Beilharz, T.H., Thorburn, D.R., et al. (2016). Accessory subunits are integral for assembly and function of human mitochondrial complex I. Nature *538*, 123–126.

Thompson, R.F., Walker, M., Siebert, C.A., Muench, S.P., and Ranson, N.A. (2016). An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. Methods *100*, 3–15.

Torraco, A., Peralta, S., Iommarini, L., and Diaz, F. (2015). Mitochondrial diseases part I: Mouse models of OXPHOS deficiencies caused by defects in respiratory complex subunits or assembly factors. Mitochondrion *21*, 76–91.

Touw, W.G., Baakman, C., Black, J., Te Beek, T.A.H., Krieger, E., Joosten, R.P., and Vriend, G. (2015). A series of PDB-related databanks for everyday needs. Nucleic Acids Res. *43*, D364–D368.

Tuschen, G., Sackmann, U., Nehls, U., Haiker, H., Buse, G., and Weiss, H. (1990). Assembly of NADH: Ubiquinone reductase (complex I) in Neurospora mitochondria. Independent pathways of nuclear-encoded and mitochondrially encoded subunits. J. Mol. Biol. *213*, 845–857.

Ugalde, C., Janssen, R.J.R.J., van den Heuvel, L.P., Smeitink, J. a M., and Nijtmans, L.G.J. (2004). Differences in assembly or stability of complex I and other mitochondrial OXPHOS complexes in inherited complex I deficiency. Hum. Mol. Genet. *13*, 659–667.

Valsecchi, F., Monge, C., Forkink, M., De Groof, A.J.C., Benard, G., Rossignol, R., Swarts, H.G., Van Emst-De Vries, S.E., Rodenburg, R.J., Calvaruso, M. a., et al. (2012). Metabolic consequences of NDUFS4 gene deletion in immortalized mouse embryonic fibroblasts. Biochim. Biophys. Acta - Bioenerg. *1817*, 1925–1936.

Valsecchi, F., Grefte, S., Roestenberg, P., Joosten-Wagenaars, J., Smeitink, J.A.M., Willems, P.H.G.M., and Koopman, W.J.H. (2013). Primary fibroblasts of NDUFS4-/- mice display increased ROS levels and aberrant mitochondrial morphology. Mitochondrion *13*, 436–443.

Varghese, F., Atcheson, E., Bridges, H.R., and Hirst, J. (2015). Characterization of clinically identified mutations in NDUFV1, the flavin-binding subunit of respiratory complex I, using a yeast model system. Hum. Mol. Genet. *24*, 6350–6360.

Verkhovskaya, M.L., Belevich, N., Euro, L., Wikström, M., and Verkhovsky, M.I. (2008). Real-time electron transfer in respiratory complex I. Proc. Natl. Acad. Sci. U. S. A. *105*, 3763–3767.

Vinothkumar, K.R., Zhu, J., and Hirst, J. (2014). Architecture of mammalian respiratory complex I. Nature *515*, 80–84.

Van Vranken, J.G., Na, U., Winge, D.R., and Rutter, J. (2015). Protein-mediated assembly of succinate dehydrogenase and its cofactors. Crit. Rev. Biochem. Mol. Biol. *50*, 168–180.
Wai, T., and Langer, T. (2016). Mitochondrial Dynamics and Metabolic Regulation. Trends Endocrinol. Metab. *27*, 105–117.

Walker, J.E., Skehel, J.M., and Buchanan, S.K. (1995). Structural analysis of NADH: ubiquinone oxidoreductase from bovine heart mitochondria. Methods Enzymol. *260*, 14–34.

Watmough, N.J., and Frerman, F.E. (2010). The electron transfer flavoprotein: Ubiquinone oxidoreductases. Biochim. Biophys. Acta - Bioenerg. *1797*, 1910–1916.

Weaver, T.M., Levitt, D.G., Donnelly, M.I., Stevens, P.P., and Banaszak, L.J. (1995). The multisubunit active site of fumarase C from *Escherichia coli*. Nat. Struct. Biol. *2*, 654–662.

Wiedemann, N., Frazier, A.E., and Pfanner, N. (2004). The Protein Import Machinery of Mitochondria. J. Biol. Chem. *279*, 14473–14476.

Wikström, M., Sharma, V., Kaila, V.R.I., Hosler, J.P., and Hummer, G. (2015). New perspectives on proton pumping in cellular respiration. Chem. Rev. *115*, 2196-2221.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2017). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci. *27*, 293-315.

Wittig, I., Braun, H.-P., and Schägger, H. (2006). Blue native PAGE. Nat. Protoc. *1*, 418–428.

Wright, J.J., Salvadori, E., Bridges, H.R., Hirst, J., and Roessler, M.M. (2016). Small-volume potentiometric titrations: EPR investigations of Fe-S cluster N2 in mitochondrial complex I. J. Inorg. Biochem. *162*, 201–206.

Wu, M., Gu, J., Guo, R., Huang, Y., and Yang, M. (2016). Structure of Mammalian Respiratory Supercomplex $I_1III_2IV_1$. Cell *167*, 1598–1609.e10.

Yip, C., Harbour, M.E., Jayawardena, K., Fearnley, I.M., and Sazanov, L.A. (2011). Evolution of Respiratory Complex I. J. Biol. Chem. *286*, 5023–5033.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zhu, J., King, M.S., Yu, M., Klipcan, L., Leslie, A.G.W., and Hirst, J. (2015). Structure of subcomplex Iβ of mammalian respiratory complex I leads to new supernumerary subunit assignments. Proc. Natl. Acad. Sci. *112*, 12087–12092.

Zhu, J., Vinothkumar, K.R., and Hirst, J. (2016). Structure of mammalian respiratory complex I. Nature *536*, 354–358.

Zhu, Z., Yao, J., Johns, T., Fu, K., De Bie, I., Macmillan, C., Cuthbert, A.P., Newbold, R.F., Wang, J.C., Chevrette, M., et al. (1998). SURF1, encoding a factor involved in the biogenesis of cytochrome c oxidase, is mutated in Leigh syndrome. Nat. Genet. *20*, 337–343.

Zickermann, V., Wirth, C., Nasiri, H., Siegmund, K., Schwalbe, H., Hunte, C., and Brandt, U. (2015). Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I. Science *347*, 44–49.

Zwicker, K., Galkin, A., Dröse, S., Grgic, L., Kerscher, S., and Brandt, U. (2006). The Redox-Bohr Group Associated with Iron-Sulfur Cluster N2 of Complex I. J. Biol. Chem. *281*, 23013–23017.